

09/889314

FILE: HCAPLUS, MEDLINE, BIOSIS, EMBASE, WPIDS, CONFSCI, SCISEARCH,
JICST-EPLUS, JAPIO' ENTERED AT 10:37:11 ON 28 JAN 2003

Author(s)

L1 584 S "BURNIE J"?/AU
L2 5349 S "MATTHEWS R"?/AU
L3 326 S L1 AND L2
L4 45 S (L3 OR L1 OR L2) AND PNEUMON?
L5 21 DUP REM L4 (24 DUPLICATES REMOVED)

L5 ANSWER 1 OF 21 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 1
ACCESSION NUMBER: 2000:553692 HCAPLUS
DOCUMENT NUMBER: 133:145931
TITLE: Protein and DNA sequences of a novel Chlamydia
pneumoniae antigen and the uses in
diagnosis and treatment of diseases associated
with Chlamydia infection
INVENTOR(S): Burnie, James Peter; Matthews,
Ruth Christine
PATENT ASSIGNEE(S): Neutec Pharma Plc, UK
SOURCE: PCT Int. Appl., 35 pp.
CODEN: PIXXD2
DOCUMENT TYPE: Patent
LANGUAGE: English
FAMILY ACC. NUM. COUNT: 1
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2000046359	A2	20000810	WO 2000-GB237	20000128
WO 2000046359	A3	20001207		
W:	AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM, AZ, BY, KG, KZ, MD, RU, TJ, TM			
RW:	GH, GM, KE, LS, MW, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG			
EP 1149162	A2	20011031	EP 2000-901235	20000128
R:	AT, BE, CH, DE, DK, ES, FR, GB, GR, IT, LI, LU, NL, SE, MC, PT, IE, FI			

PRIORITY APPLN. INFO.: GB 1999-2555 A 19990205
WO 2000-GB237 W 20000128

AB The invention provides protein and DNA sequences of a novel Chlamydia **pneumoniae** antigen. The present invention further relates to the uses of the antigens of this invention in treatment, prevention and diagnosis of infection due to Chlamydia **pneumoniae** and in particular to the prevention and treatment of atherosclerosis, including coronary atherosclerosis, caused by same.

L5 ANSWER 2 OF 21 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 1999:126671 BIOSIS
DOCUMENT NUMBER: PREV199900126671
TITLE: Is infection control an academic study.
AUTHOR(S): Burnie, J. P. (1)
CORPORATE SOURCE: (1) Dep. Med. Microbiol., 2nd Floor, Clin. Sci.
Build., Manchester Royal Infirmary, Oxford Road,

THIS PAGE BLANK (USPTO)

09/889314

SOURCE: Manchester M13 9WL UK
Journal of Hospital Infection, (Jan., 1999) Vol. 41,
No. 1, pp. 7-10.
ISSN: 0195-6701.

DOCUMENT TYPE: Editorial

LANGUAGE: English

AB This editorial aims to answer the question of whether infection control is an academic specialty. By considering the consequences of a lack of infection control in terms of patient morbidity and mortality and hence cost, it is easy to establish the importance of the area. Infection control embraces not only developing policies for preventing the physical spread of a microorganism but also prophylactic therapy such as vaccination and therapeutic measures such as antibiotics. Infection control not only applies to localized infection in hospital due to antibiotic resistant microbes but also to the community. Bacteria such as *Helicobacter pylori* and *Chlamydia pneumoniae* and the viruses Hepatitis B, Hepatitis C, human lymphotropic virus type 1, Epstein-Barr viruses and human papilloma virus have been implicated in diseases not previously thought to have an infectious origin. Coping with these problems is clearly an academic area.

L5 ANSWER 3 OF 21 MEDLINE DUPLICATE 2

ACCESSION NUMBER: 2000028548 MEDLINE
DOCUMENT NUMBER: 20028548 PubMed ID: 10562808
TITLE: A polymerase chain reaction enzyme immunoassay for diagnosing infection caused by *Aspergillus fumigatus*.
AUTHOR: Golbang N; Burnie J P; Matthews R C
CORPORATE SOURCE: Department of Medical Microbiology, Manchester University, Manchester Royal Infirmary, UK.
SOURCE: JOURNAL OF CLINICAL PATHOLOGY, (1999 Jun) 52 (6) 419-23.
Journal code: 0376601. ISSN: 0021-9746.
PUB. COUNTRY: ENGLAND: United Kingdom
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Abridged Index Medicus Journals; Priority Journals
ENTRY MONTH: 199911
ENTRY DATE: Entered STN: 20000111
Last Updated on STN: 20000111
Entered Medline: 19991116

AB AIM: To develop a polymerase chain reaction enzyme immunoassay (PCR-EIA) to measure levels of circulating aspergillus DNA in invasive aspergillosis caused by *Aspergillus fumigatus*. METHODS: The PCR reaction was based on primers from the 18s rRNA gene. Binding of the product to a streptavidin coated microtitration plate was mediated by a biotinylated capture probe. The product was digoxigenylated during PCR and this was the tag to which antibody was bound in the subsequent EIA. RESULTS: The optical density (OD) endpoint was < 0.1 in 10 sera from neutropenic patients with no evidence of invasive aspergillosis, and in 10 sera from nonneutropenic patients with bacterial pneumonia (group 1). The OD from five of 12 patients with allergic bronchopulmonary aspergillosis (ABPA) (group 2), three with an aspergilloma (group 3), and five with possible invasive aspergillosis (group 4) was > or = 0.1. In 63 sera from 33 cases of proven invasive aspergillosis (group 5) an OD > or = 0.1 was achieved in 48 sera from 30 patients. The maximum OD was 0.510. The level fell in survivors and gradually

THIS PAGE BLANK (USPTO)

09/889314

rose in fatal cases. CONCLUSIONS: This assay validated the concept of diagnosing invasive aspergillosis by measuring levels of circulating fungal DNA in serum.

L5 ANSWER 4 OF 21 MEDLINE DUPLICATE 3
ACCESSION NUMBER: 1999135117 MEDLINE
DOCUMENT NUMBER: 99135117 PubMed ID: 9949958
TITLE: Is infection control an academic study?.
AUTHOR: **Burnie J P**
CORPORATE SOURCE: Department of Medical Microbiology, Manchester Royal Infirmary.
SOURCE: JOURNAL OF HOSPITAL INFECTION, (1999 Jan) 41 (1) 7-10. Ref: 30
Journal code: 8007166. ISSN: 0195-6701.
PUB. COUNTRY: ENGLAND: United Kingdom
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
General Review; (REVIEW)
(REVIEW, TUTORIAL)
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199904
ENTRY DATE: Entered STN: 19990426
Last Updated on STN: 19990426
Entered Medline: 19990413

AB This editorial aims to answer the question of whether infection control is an academic specialty. By considering the consequences of a lack of infection control in terms of patient morbidity and mortality and hence cost, it is easy to establish the importance of the area. Infection control embraces not only developing policies for preventing the physical spread of a micro-organism but also prophylactic therapy such as vaccination and therapeutic measures such as antibiotics. Infection control not only applies to localized infection in hospital due to antibiotic resistant microbes but also to the community. Bacteria such as *Helicobacter pylori* and *Chlamydia pneumoniae* and the viruses hepatitis B, hepatitis C, human lymphotropic virus type 1, Epstein-Barr viruses and human papilloma virus have been implicated in diseases not previously thought to have an infectious origin. Coping with these problems is clearly an academic area.

L5 ANSWER 5 OF 21 EMBASE COPYRIGHT 2003 ELSEVIER SCI. B.V.
ACCESSION NUMBER: 97018133 EMBASE
DOCUMENT NUMBER: 1997018133
TITLE: Cervical cancer as an AIDS-defining illness.
AUTHOR: Maiman M.; Fruchter R.G.; Clark M.; Arrastia C.D.;
Matthews R.; Gates E.J.
CORPORATE SOURCE: Dr. M. Maiman, Division of Gynecologic Oncology,
Department of Obstetrics/Gynecology, SUNY-Health
Science Center, 450 Clarkson Avenue, Brooklyn, NY
11203, United States
SOURCE: Obstetrics and Gynecology, (1997) 89/1 (76-80).
Refs: 19
ISSN: 0029-7844 CODEN: OBGNAS
PUBLISHER IDENT.: S 0029-7844(96)00378-X
COUNTRY: United States
DOCUMENT TYPE: Journal; Article
FILE SEGMENT: 004 Microbiology
010 Obstetrics and Gynecology

Searcher : Shears 308-4994

THIS PAGE BLANK (USPTO)

016 Cancer
 017 Public Health, Social Medicine and
 Epidemiology

LANGUAGE: English

SUMMARY LANGUAGE: English

AB Objective: To evaluate the importance of cervical cancer in the spectrum of human immunodeficiency virus (HIV) related diseases at a single high-risk institution and to compare disease characteristics in HIV-infected women with cervical cancer and those with other AIDS-related malignancies. Methods: We retrospectively reviewed data on cervical cancer and AIDS in women registered through the New York City Department of Health and institutional tumor registries from 1987 through 1995. Results: During the study period, cervical cancer was diagnosed in 28 HIV-positive women. In 26, cervical cancer was the initial AIDS-defining illness, representing 4% (26 of 725) of the subjects, and it was the sixth most common initial AIDS-defining illness in women. Cervical cancer was the most common AIDS-related malignancy among women, representing 55% of the cases, followed by lymphoma (29%) and Kaposi sarcoma (16%). In 71% of the women with cervical cancer, HIV infection was diagnosed at the time of cancer presentation by routine testing, whereas in women with other malignancies, HIV diagnosis preceded cancer diagnosis (70%) by a mean of 2.7 years. Patients with other malignancies had greater immunosuppression (mean CD4 count 153/.mu.L) than those with cervical cancer (mean CD4 count 312/.mu.L). The recurrence rate for women with cervical cancer was 88%. Although the interval from cancer diagnosis to death was similar in all three groups (9.1-12.4 months), cancer was the cause of death in 95% of HIV-infected women with cervical cancer, compared with 60% of those with other AIDS-related malignancies. Conclusion: In urban populations at increased risk for both diseases, cervical cancer is an important AIDS-defining illness and may be the most common AIDS-related malignancy in women.

L5 ANSWER 6 OF 21 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.

ACCESSION NUMBER: 1996:109315 BIOSIS

DOCUMENT NUMBER: PREV199698681450

TITLE: Immunoblotting in the diagnosis of culture negative endocarditis caused by streptococci and enterococci.

AUTHOR(S): Burnie, J. P. (1); Clark, I.

CORPORATE SOURCE: (1) Dep. Medical Microbiol., 2nd Floor, Clinical Sci. Build., Manchester Royal Infirmary, Oxford Road, Manchester M13 9WL UK

SOURCE: Journal of Clinical Pathology (London), (1995) Vol. 48, No. 12, pp. 1130-1136.
 ISSN: 0021-9746.

DOCUMENT TYPE: Article

LANGUAGE: English

AB Aim: To improve the diagnosis of culture negative endocarditis by diagnosing cases due to streptococci and enterococci. Method: Serum samples were immunoblotted against extracts of the commonest streptococci and enterococci. They were selected from patients with a cardiac murmur, persistent pyrexia and at least three negative blood cultures. The presence of patterns of endocarditis species specific antigenic bands was measured and correlated with clinical outcome. Results: Negative serology was found in 28 patients where the diagnosis of endocarditis was rejected or, if proved, staphylococcal, yeast, Gram negative, systemic lupus erythematosus,

THIS PAGE BLANK (USPTO)

due to Q fever or Chlamydia psittaci or nonbacterial thrombotic. Positive serology was found in 27 of the 34 patients where the response to antibiotics suggested streptococcal or enterococcal infection. In 22 of these there was objective evidence of endocarditis. Positive serology was also found in three of four further patients with vegetations at necropsy. Conclusion: The identification of patterns of antibody response on immunoblotting can be used to make a specific diagnosis of streptococcal or enterococcal endocarditis in the absence of positive blood cultures.

L5 ANSWER 7 OF 21 SCISEARCH COPYRIGHT 2003 ISI (R)
 ACCESSION NUMBER: 94:539953 SCISEARCH
 THE GENUINE ARTICLE: PE730
 TITLE: THE LEUCINE-RESPONSIVE REGULATORY PROTEIN, A GLOBAL REGULATOR OF METABOLISM IN ESCHERICHIA-COLI
 AUTHOR: CALVO J M (Reprint); MATTHEWS R G
 CORPORATE SOURCE: CORNELL UNIV, BIOCHEM MOLEC & CELL BIOL SECT, ITHACA, NY, 14853 (Reprint); UNIV MICHIGAN, DIV BIOPHYS RES, ANN ARBOR, MI, 48109; UNIV MICHIGAN, DEPT BIOL CHEM, ANN ARBOR, MI, 48109
 COUNTRY OF AUTHOR: USA
 SOURCE: MICROBIOLOGICAL REVIEWS, (SEP 1994) Vol. 58, No. 3, pp. 466-490.
 ISSN: 0146-0749.
 DOCUMENT TYPE: General Review; Journal
 FILE SEGMENT: LIFE
 LANGUAGE: ENGLISH
 REFERENCE COUNT: 187

ABSTRACT IS AVAILABLE IN THE ALL AND IALL FORMATS

AB Natural genetic transformation is the active uptake of free DNA by bacterial cells and the heritable incorporation of its genetic information. Since the famous discovery of transformation in *Streptococcus pneumoniae* by Griffith in 1928 and the demonstration of DNA as the transforming principle by Avery and coworkers in 1944, cellular processes involved in transformation have been studied extensively by in vitro experimentation with a few transformable species. Only more recently has it been considered that transformation may be a powerful mechanism of horizontal gene transfer in natural bacterial populations. In this review the current understanding of the biology of transformation is summarized to provide the platform on which aspects of bacterial transformation in water soil, and sediments and the habitat of pathogens are discussed. Direct and indirect evidence for gene transfer routes by transformation within species and between different species will be presented, along with data suggesting that plasmids as well as chromosomal DNA are subject to genetic exchange via transformation. Experiments exploring the prerequisites for transformation in the environment, including the production and persistence of free DNA and factors important for the uptake of DNA by cells, will be compiled, as well as possible natural barriers to transformation. The efficiency of gene transfer by transformation in bacterial habitats is possibly genetically adjusted to submaximal levels. The fact that natural transformation has been detected among bacteria from all trophic and taxonomic groups including archaeobacteria suggests that transformability evolved early in phylogeny. Probable functions of DNA uptake other than gene acquisition will be discussed. The body of information presently available suggests that transformation has a great impact on bacterial population dynamics

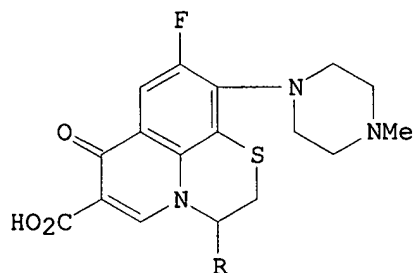
THIS PAGE BLANK (USPTO)

09/889314

as well as on bacterial evolution and speciation.

L5 ANSWER 8 OF 21 SCISEARCH COPYRIGHT 2003 ISI (R)
ACCESSION NUMBER: 92:717086 SCISEARCH
THE GENUINE ARTICLE: KA952
TITLE: PROGRESSIVE CHRONIC PULMONARY ASPERGILLOSIS - A
DIAGNOSTIC AND THERAPEUTIC CHALLENGE
AUTHOR: ELLIS M E (Reprint); DOSSING M; ALHOKAIL A; QADRI S
H; HAINAU B; **BURNIE J**
CORPORATE SOURCE: KING FAISAL SPECIALIST HOSP & RES CTR, DEPT MED, MBC
46, POB 3354, RIYADH 11211, SAUDI ARABIA (Reprint);
KING FAISAL SPECIALIST HOSP & RES CTR, DEPT PATHOL &
LAB MED, RIYADH 11211, SAUDI ARABIA; UNIV
MANCHESTER, SCH MED, DEPT MICROBIOL, MANCHESTER M13
9PL, LANCs, ENGLAND
COUNTRY OF AUTHOR: SAUDI ARABIA; ENGLAND
SOURCE: JOURNAL OF THE ROYAL SOCIETY OF MEDICINE, (DEC 1992)
Vol. 85, No. 12, pp. 763-764.
ISSN: 0141-0768.
DOCUMENT TYPE: Article; Journal
FILE SEGMENT: LIFE; CLIN
LANGUAGE: ENGLISH
REFERENCE COUNT: 13

L5 ANSWER 9 OF 21 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 4
ACCESSION NUMBER: 1992:507995 HCAPLUS
DOCUMENT NUMBER: 117:107995
TITLE: The in vitro activity of two new quinolones:
rufloxacin and MF 961
AUTHOR(S): Wise, R.; Andrews, J. M.; **Matthews, R.**
; Wolstenholme, M.
CORPORATE SOURCE: Dep. Med. Microbiol., Dudley Road Hosp.,
Birmingham, B18 7QH, UK
SOURCE: Journal of Antimicrobial Chemotherapy (1992),
29(6), 649-60
CODEN: JACHDX; ISSN: 0305-7453
DOCUMENT TYPE: Journal
LANGUAGE: English
GI



I, R=H

II, R=CH₂F

AB The in vitro activity of two new quinolone antimicrobials,
rufloxacin (I) and MF 961 (II), together with the desmethylated
metabolite of rufloxacin (MF 922) were compared with other orally
administered agents against 622 bacterial strains. Against

THIS PAGE BLANK (USPTO)

09/889314

Enterobacteriaceae and *Pseudomonas aeruginosa*, I was generally active (MIC90 1-8 mg/L) with the exception of *Klebsiella* and *Serratia* spp. (MIC90 32 mg/L) and *Enterobacter* spp. (MIC90 64 mg/L). The respiratory pathogens *Haemophilus influenzae* and *Moraxella catarrhalis* were susceptible to I (MIC90 0.5 and 1 mg/L, resp.), but *Streptococcus pneumoniae* was less susceptible (MIC90 32 mg/L). *Staphylococcus aureus* was susceptible to I (MIC90 2 mg/L). The I metabolite, MF 922, was generally as active as its parent. II was usually 2-fold more active than I. All 3 compds. were 4-16 times less active than norfloxacin, but I was as active or somewhat more active than norfloxacin against *Staphylococcus* spp. Any strains showing decreased susceptibility to other quinolones exhibited cross resistance to these new agents. The MBC of I and MF 922 was within one diln. of the MIC, and human serum had little effect upon the activity of both agents. The protein binding of I and MF 922 at 1 and 10 mg/L was 55 and 63.8% and 30.3% and 32.6%, resp. The activity of rifloxacin against four strains of *Chlamydia trachomatis* and one strain of *Chlamydia pneumoniae* was detd. The MIC for *C. trachomatis* was 4-8 mg/L and 4 mg/L for *C. pneumoniae*.

L5 ANSWER 10 OF 21 MEDLINE
ACCESSION NUMBER: 92115389 MEDLINE
DOCUMENT NUMBER: 92115389 PubMed ID: 1766717
TITLE: *Chlamydia pneumoniae* (TWAR) in neonates.
AUTHOR: Matthews R S; Mohite A; Addy D P; Wise R
SOURCE: PEDIATRIC INFECTIOUS DISEASE JOURNAL, (1991 Dec) 10
(12) 956-7.
Journal code: 8701858. ISSN: 0891-3668.
PUB. COUNTRY: United States
DOCUMENT TYPE: Letter
LANGUAGE: English
FILE SEGMENT: Priority Journals
ENTRY MONTH: 199202
ENTRY DATE: Entered STN: 19920308
Last Updated on STN: 19920308
Entered Medline: 19920214

L5 ANSWER 11 OF 21 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
DUPLICATE 5
ACCESSION NUMBER: 1991:226713 BIOSIS
DOCUMENT NUMBER: BA91:118173
TITLE: ISOLATION CLONING MAPPING AND NUCLEOTIDE SEQUENCING
OF THE GENE ENCODING FLAVODOXIN IN *ESCHERICHIA-COLI*.
AUTHOR(S): OSBORNE C; CHEN L-M; MATTHEWS R G
CORPORATE SOURCE: DEP. BIOL. CHEM., UNIV. MICHIGAN, ANN ARBOR, MICH.
48109.
SOURCE: J BACTERIOL, (1991) 173 (5), 1729-1737.
CODEN: JOBAAY. ISSN: 0021-9193.
FILE SEGMENT: BA; OLD
LANGUAGE: English

AB The flavodoxins constitute a highly conserved family of small, acidic electron transfer proteins with flavin mononucleotide prosthetic groups. They are found in prokaryotes and in red and green algae, where they provide electron at low potentials for the reduction of nitrogen by nitrogenase, for the light-dependent reduction of NADP+ in photosynthesis, and for the reduction of sulfite. Proteins with the physical characteristics of flavodoxins

THIS PAGE BLANK (USPTO)

have been implicated in the reductive activation of pyruvate formate-lyase and cobalamin-dependent methionine synthase in *Escherichia coli*. We have purified flavodoxin to homogeneity from *E. coli*, determined its N-terminal amino acid sequence, and used this sequence to construct a 64-fold degenerate oligonucleotide probe for the flavodoxin gene. Because the phenotype of a flavodoxin mutant is not known, we used this degenerate probe to screen the phages of Kohara library and identified two phages, with inserts mapping at .apprx. 16 min, that hybridized to the probe. The flavodoxin gene, designated *fldA*, was subcloned from the DNA in the overlap region of these two clones. The deduced amino acid sequence, determined by nucleotide sequencing of the flavodoxin gene, shows strong homology with flavodoxins from nitrogen-fixing bacteria and cyanobacteria. The *fldA* gene maps at 15.9 min on the *E. coli* chromosome and is transcribed in a counterclockwise direction.

L5 ANSWER 12 OF 21 EMBASE COPYRIGHT 2003 ELSEVIER SCI. B.V.DUPLICATE
6

ACCESSION NUMBER: 92041225 EMBASE
DOCUMENT NUMBER: 1992041225
TITLE: Chlamydia **pneumoniae** (TWAR) in neonates
[1].
AUTHOR: **Matthews R.S.**; Mohite A.; Addy D.P.; Wise
R.
CORPORATE SOURCE: Dudley Road Hospital, Birmingham, United Kingdom
SOURCE: Pediatric Infectious Disease Journal, (1991) 10/12
(956-957).
ISSN: 0891-3668 CODEN: PIDJEV
COUNTRY: United States
DOCUMENT TYPE: Journal; Letter
FILE SEGMENT: 004 Microbiology
007 Pediatrics and Pediatric Surgery
030 Pharmacology
037 Drug Literature Index
LANGUAGE: English

L5 ANSWER 13 OF 21 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 7
ACCESSION NUMBER: 1991:651968 HCAPLUS
DOCUMENT NUMBER: 115:251968
TITLE: In-vitro susceptibility of Chlamydia
pneumoniae (TWAR) to seven antibiotics
AUTHOR(S): Cooper, M. A.; Baldwin, D.; **Matthews, R.**
S.; Andrews, J. M.; Wise, R.
CORPORATE SOURCE: Dep. Microbiol., Dudley Road Hosp., Birmingham,
UK
SOURCE: Journal of Antimicrobial Chemotherapy (1991),
28(3), 407-13
CODEN: JACHDX; ISSN: 0305-7453
DOCUMENT TYPE: Journal
LANGUAGE: English

AB A modification of an immunofluorescence method previously used to study the in vitro antimicrobial susceptibilities of *C. trachomatis* was used to investigate the activity of seven antimicrobials against a strain of *C. pneumoniae*. Adding antimicrobial to preinfected cells gave higher MICs (min. inhibitory concns.) and MLCs (min. lethal concns.) than when cells were infected in the presence of the antimicrobials, and this difference in methodol. could account for the discrepancy between these results and those of

THIS PAGE BLANK (USPTO)

09/889314

others. Of the antimicrobials studied, clarithromycin and its 14-hydroxy metabolite were the most active agents; sparfloxacin was more active than ciprofloxacin, but no more active than more conventional antichlamydial agents.

L5 ANSWER 14 OF 21 BIOSIS COPYRIGHT 2003 BIOLOGICAL ABSTRACTS INC.
ACCESSION NUMBER: 1991:535104 BIOSIS
DOCUMENT NUMBER: BR41:124839
TITLE: THE ACTIVITY OF SPARFLOXACIN AND OTHER AGENTS AGAINST CHLAMYDIA-**PNEUMONIAE** THE EFFECT OF PRE AND POST ANTIBIOTIC EXPOSURE.
AUTHOR(S): WISE R; ANDREWS J M; COOPER M A; **MATTHEWS R**
CORPORATE SOURCE: DUDLEY ROAD HOSP., BIRMINGHAM B18 7QH, UK.
SOURCE: THIRTY-FIRST INTERSCIENCE CONFERENCE ON ANTIMICROBIAL AGENTS AND CHEMOTHERAPY, CHICAGO, ILLINOIS, USA, SEPTEMBER 29-OCTOBER 2, 1991. PROGRAM ABSTR, (1991) 31 (0), 212.
CODEN: PCHES.
DOCUMENT TYPE: Conference
FILE SEGMENT: BR; OLD
LANGUAGE: English

L5 ANSWER 15 OF 21 MEDLINE DUPLICATE 8
ACCESSION NUMBER: 91324522 MEDLINE
DOCUMENT NUMBER: 91324522 PubMed ID: 1907617
TITLE: Immunoblotting and culture positive endocarditis.
AUTHOR: Clark I; **Burnie J P**
CORPORATE SOURCE: Department of Medical Microbiology, University of Manchester Medical School.
SOURCE: JOURNAL OF CLINICAL PATHOLOGY, (1991 Feb) 44 (2) 152-6.
Journal code: 0376601. ISSN: 0021-9746.
PUB. COUNTRY: ENGLAND: United Kingdom
DOCUMENT TYPE: Journal; Article; (JOURNAL ARTICLE)
LANGUAGE: English
FILE SEGMENT: Abridged Index Medicus Journals; Priority Journals
ENTRY MONTH: 199109
ENTRY DATE: Entered STN: 19910929
Last Updated on STN: 19910929
Entered Medline: 19910909

AB Serum samples from patients with endocarditis due to Streptococcus mutans, Streptococcus **pneumoniae**, Streptococcus agalactiae, Streptococcus lactis and a "nutritionally" variant streptococcus were immunoblotted against antigenic extracts from all five species. In S mutans endocarditis there was an endocarditis specific pattern of IgM against bands of 220, 200, and 190 kilodaltons. In S **pneumoniae** IgM against antibody of a molecular weight greater than 150 kilodaltons was specific to endocarditis. In S agalactiae IgM against bands at 82, 71, and 66-67 kilodaltons was endocarditis specific. In S lactis endocarditis specific IgM was present against antigenic bands at 105, 66, 61 and 58 kilodaltons. With the "nutritionally" variant streptococcus it was impossible to distinguish between cases of endocarditis and controls.

L5 ANSWER 16 OF 21 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 9
ACCESSION NUMBER: 1991:58770 HCAPLUS
DOCUMENT NUMBER: 114:58770

Searcher : Shears 308-4994

THIS PAGE BLANK (USPTO)

09/889314

TITLE: In vitro activity of sparfloxacin, a new
quinolone antimicrobial agent
AUTHOR(S): Cooper, M. A.; Andrews, J. M.; Ashby, J. P.;
Matthews, R. S.; Wise, R.
CORPORATE SOURCE: Dep. Microbiol., Dudley Road Hosp., Birmingham,
UK
SOURCE: Journal of Antimicrobial Chemotherapy (1990),
26(5), 667-76
CODEN: JACHDX; ISSN: 0305-7453
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The in vitro activity of sparfloxacin (AT-4140), a new difluorinated quinolone, was compared with those of ciprofloxacin, temafloxacin, and selected members of other groups of antimicrobial agents, against 651 recent distinct clin. isolates and strains with known mechanisms of resistance. Three strains of *Chlamydia trachomatis* were also studied. The MICs for 90% of the Enterobacteriaceae were 0.06-1 mg/L; for *Pseudomonas aeruginosa* the MIC₉₀ was 2 mg/L. Sparfloxacin was 16-fold more active against *Acinetobacter* spp. than ciprofloxacin. For *Staphylococcus* spp., *Streptococcus* spp., and *Enterococcus faecalis*, the MIC₉₀ was 0.25-1 mg/L; sparfloxacin was 4-fold more active against *Streptococcus pneumoniae* than ciprofloxacin. Ninety percent of strains of *Haemophilus influenzae*, *Branhamella catarrhalis*, and *Neisseria* spp. were inhibited by <0.03 mg/L; for *Bacteroides fragilis* the MIC₉₀ was 1 mg/L. The 3 strains of *C. trachomatis* were susceptible to 0.06-0.12 mg/L sparfloxacin, which was 16-fold more active than ciprofloxacin. There was cross resistance among the quinolones, but not between the quinolones and other groups of antimicrobials. The protein binding of sparfloxacin was 40% and serum had little effect on its activity.

L5 ANSWER 17 OF 21 HCAPLUS COPYRIGHT 2003 ACS DUPLICATE 10

ACCESSION NUMBER: 1988:435176 HCAPLUS
DOCUMENT NUMBER: 109:35176
TITLE: In vitro activity of lomefloxacin, a new
quinolone antimicrobial agent, in comparison
with those for other agents
AUTHOR(S): Wise, R.; Andrews, J. M.; Ashby, J. P.;
Matthews, R. S.
CORPORATE SOURCE: Dep. Med. Microbiol., Dudley Road Hosp.,
Birmingham, B18 7QH, UK
SOURCE: Antimicrobial Agents and Chemotherapy (1988),
32(5), 617-22
CODEN: AMACQ; ISSN: 0066-4804
DOCUMENT TYPE: Journal
LANGUAGE: English

AB The in vitro activity of lomefloxacin (I), a new difluorinated quinolone, was compared with those of ofloxacin, ciprofloxacin, fleroxacin, amoxicillin, cefuroxime, and trimethoprim against 585 recent clin. isolates and other strains with known mechanisms of resistance. The MICs of I against 90% of the members of the family Enterobacteriaceae, *Pseudomonas aeruginosa*, and staphylococci were between 0.25 and 4 .mu.g/mL. Ninety percent of *Neisseria* sp. and *Haemophilus influenzae* were susceptible to .ltoreq.0.06 .mu.g/mL, and streptococci (including *S. pyogenes*, *S. pneumoniae*, and enterococci) and *Bacteroides feragilis* were susceptible to 8 .mu.g/mL. I was comparable in activity to fleroxacin and ofloxacin, but it was less active than ciprofloxacin. There was

THIS PAGE BLANK (USPTO)

09/889314

cross-resistance between the quinolone group of antimicrobial agents. The protein binding of I was 15.4%, and serum had little effect on the activity of the compd. However, urine at pH 5.0 decreased the activity by 2-8-fold compared with that at pH 7.0.

L5 ANSWER 18 OF 21 EMBASE COPYRIGHT 2003 ELSEVIER SCI. B.V.

ACCESSION NUMBER: 88049469 EMBASE

DOCUMENT NUMBER: 1988049469

TITLE: Chlamydial rhinitis neonatorum.

AUTHOR: Kent S.E.; **Matthews R.S.**

CORPORATE SOURCE: Birmingham Children's Hospital, Ladywood, Birmingham B16 8ET, United Kingdom

SOURCE: Journal of Laryngology and Otology, (1987) 101/11 (1193-1197).

ISSN: 0022-2151 CODEN: JLOTAX

COUNTRY: United Kingdom

DOCUMENT TYPE: Journal

FILE SEGMENT: 004 Microbiology
007 Pediatrics and Pediatric Surgery
010 Obstetrics and Gynecology
011 Otorhinolaryngology

LANGUAGE: English

SUMMARY LANGUAGE: English

AB This paper reports the case of a five-week-old child who developed severe rhinitis with epistaxes, due to infection with chlamydia trachomatis. The infection was acquired from the mother's genital tract during birth but the true diagnosis was not suspected for some time. Chlamydia trachomatis is well recognized as one of the causes of neonatal conjunctivitis and **pneumonia**. However, its presence in the upper respiratory tract is thought usually to be asymptomatic (Oriol and Ridgway, 1982). Little attention has been focused on its role in neonatal rhinitis. This paper draws to the otolaryngologist's attention the entity of chlamydial rhinitis neonatorum and discusses the epidemiology, diagnosis and management of the condition.

L5 ANSWER 19 OF 21 EMBASE COPYRIGHT 2003 ELSEVIER SCI. B.V.

ACCESSION NUMBER: 85172703 EMBASE

DOCUMENT NUMBER: 1985172703

TITLE: CM-40874.

AUTHOR: **Burnie J.; Matthews R.**

CORPORATE SOURCE: United Kingdom

SOURCE: Drugs of the Future, (1985) 10/3 (193-195).

CODEN: DRFUD4

COUNTRY: Spain

DOCUMENT TYPE: Journal

FILE SEGMENT: 037 Drug Literature Index
052 Toxicology

LANGUAGE: English

L5 ANSWER 20 OF 21 MEDLINE

ACCESSION NUMBER: 78198713 MEDLINE

DOCUMENT NUMBER: 78198713 PubMed ID: 78425

TITLE: Multiply resistant pneumococcus.

AUTHOR: Meers P D; **Matthews R B**

SOURCE: LANCET, (1978 Jul 22) 2 (8082) 219.

Journal code: 2985213R. ISSN: 0140-6736.

PUB. COUNTRY: ENGLAND: United Kingdom

THIS PAGE BLANK (USPTO)

09/889314

DOCUMENT TYPE: Letter
LANGUAGE: English
FILE SEGMENT: Abridged Index Medicus Journals; Priority Journals
ENTRY MONTH: 197809
ENTRY DATE: Entered STN: 19900314
Last Updated on STN: 19950206
Entered Medline: 19780901

L5 ANSWER 21 OF 21 EMBASE COPYRIGHT 2003 ELSEVIER SCI. B.V.
ACCESSION NUMBER: 78377563 EMBASE
DOCUMENT NUMBER: 1978377563
TITLE: Multiply resistant pneumococcus.
AUTHOR: Meers P.D.; **Matthews R.B.**
CORPORATE SOURCE: Publ. Hlth Lab., Gen. Hosp., Plymouth PL4 8NN, United Kingdom
SOURCE: Lancet, (1978) 2/8081 (155-156).
CODEN: LANCAO
COUNTRY: United Kingdom
DOCUMENT TYPE: Journal
FILE SEGMENT: 037 Drug Literature Index
004 Microbiology
LANGUAGE: English

AB A pneumococcus resistant to chloramphenicol, tetracycline and penicillin was isolated from a holidaymaker from Spain, who was admitted to hospital in Plymouth with an acute skin condition. She had a productive cough, and from mucopurulent sputum a pneumococcus (type 23, Danish nomenclature) was isolated. On routine testing this was found to be resistant to chloramphenicol and tetracycline, and the zone of inhibition round the penicillin disc (1.5 units) was reduced from the expected .gtoreq.10 mm to 7 mm, measured radially from the edge of the disc. By the agar dilution method, using a minimal-growth endpoint, the minimum inhibitory concentration (M.I.C.) of penicillin was 0.25 .mu.g/ml compared with an M.I.C. of 0.03 .mu.g/ml for a sensitive strain. The M.I.C.s of tetracycline and chloramphenicol were both 20 .mu.g/ml. The inoculum used in these tests was about 106 organisms.

FILE 'HOME' ENTERED AT 10:38:23 ON 28 JAN 2003

19 18 17 16 15

THIS PAGE BLANK (USPTO)

977 6 1.2 515 1 US-08-063-552-4 Sequence 4, Appli
 978 6 1.2 515 5 PCT-US93-05704-4 Patent No. RE34,60
 979 6 1.2 522 6 RE34606-6 Sequence 2, Appli
 980 6 1.2 525 1 US-08-348-891A-2 Sequence 2, Appli
 981 6 1.2 525 2 US-08-905-817-2 Sequence 8, Appli
 982 6 1.2 525 2 US-09-351-224E-8 Sequence 40, Appli
 983 6 1.2 526 2 US-08-853-659A-40 Sequence 2, Appli
 984 6 1.2 529 1 US-08-548-509-2 Sequence 88, Appli
 985 6 1.2 532 4 US-09-071-035-88 Sequence 49, Appli
 986 6 1.2 535 1 US-07-872-644-49 Sequence 49, Appli
 987 6 1.2 535 1 US-08-297-510-49 Sequence 49, Appli
 988 6 1.2 535 1 US-08-455-526-49 Sequence 49, Appli
 989 6 1.2 535 1 US-08-455-526-49 Sequence 49, Appli
 990 6 1.2 535 1 US-08-455-526-49 Sequence 49, Appli
 991 6 1.2 535 1 US-08-455-526-49 Sequence 49, Appli
 992 6 1.2 535 3 US-08-813-574-2 Sequence 2, Appli
 993 6 1.2 535 3 US-09-139-491-49 Sequence 49, Appli
 994 6 1.2 535 5 PCT-US92-03222-49 Sequence 49, Appli
 995 6 1.2 540 4 US-09-011-074-4 Sequence 17, Appli
 996 6 1.2 542 4 US-09-380-420C-17 Sequence 3, Appli
 997 6 1.2 545 4 US-09-351-224E-3 Sequence 2, Appli
 998 6 1.2 550 1 US-08-484-493-2 Sequence 2, Appli
 999 6 1.2 550 2 US-08-484-494-2 Sequence 2, Appli
 1000 6 1.2 550 2 US-08-345-212-2

ALIGNMENTS

102 (e)

RESULT 1
 US-08-809-326A-1
 Sequence 1, Application US/08809326A

GENERAL INFORMATION:
 PATENT NO. 6165478
 APPLICANT: IZUTSU, HIROSHI
 APPLICANT: OBARA, KAZUHIKO
 APPLICANT: MATSUMOTO, AKIRA
 TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,
 TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING
 TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMANTS
 TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF
 TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT
 TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA
 TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND/OR
 TITLE OF INVENTION: MEASUREMENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
 TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
 NUMBER OF SEQUENCES: 31
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Pennie & Edmonds LLP
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036-2711

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/809,326A
 FILING DATE: 19-MAR-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 224711/94
 FILING DATE: 20-SEP-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 106006/95
 FILING DATE: 28-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 106008/95
 FILING DATE: 28-APR-1995

PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 106009/95
 FILING DATE: 28-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 106010/95
 FILING DATE: 28-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 106011/95
 FILING DATE: 28-APR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Miller, Charles E.
 REGISTRATION NUMBER: 24,576
 REFERENCE/DOCKET NUMBER: 7426-043-999
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864/9741
 TELETYPE: (212) 869-8864/9741
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 488 amino acids
 TYPE: amino acid
 TOPOLOGY: linear
 MOLECULE TYPE: peptide
 US-08-809-326A-1

Query Match 98.2%; Score 487; DB 4; Length 488;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MSISSSGPDNOKNIMSOVLSTPGVPOODKLSGNETKOIOFROGKNTMESDATTG 63
 DB 1 MSISSSGPDNOKNIMSOVLSTPGVPOODKLSGNETKOIOFROGKNTMESDATTG 60
 QY 64 ASGKDKTSTTETETAPPOGVAAGKESSESOKAGADTVSGAAATTAATATKIAMQTSI 123
 DB 61 ASGKDKTSTTETETAPPOGVAAGKESSESOKAGADTVSGAAATTAATATKIAMQTSI 120
 QY 124 EASMSMESTLESLOSLSAOKMEVEAYVVAALSGKSSGSAKLEPPEPKPQVPRSEVI 183
 DB 121 EASMSMESTLESLOSLSAOKMEVEAYVVAALSGKSSGSAKLEPPEPKPQVPRSEVI 180
 QY 184 EIGLALRAIOTLGATSAISLNVASTOQADOTKLGLEKOAIRIDEREYQEMKAAE 243
 DB 181 EIGLALRAIOTLGATSAISLNVASTOQADOTKLGLEKOAIRIDEREYQEMKAAE 240
 QY 244 OKSKDEGTMDTVNTVMIAVSAITVISIAAIFTCGAGLAGLAAGAAVGAAGAAGA 303
 DB 241 OKSKDEGTMDTVNTVMIAVSAITVISIAAIFTCGAGLAGLAAGAAVGAAGAAGA 300
 QY 304 AATTVAAQITVQAVVQAVKQAVITAVROATTAATKAIVKSGIKAFITVYATKAISK 363
 DB 301 AATTVAAQITVQAVVQAVKQAVITAVROATTAATKAIVKSGIKAFITVYATKAISK 360
 QY 364 GISKVFAGKTQIAKNFPLSKVSISSKVTYVGVVVAAPALGKGIQMOQLEMOON 423
 DB 361 GISKVFAGKTQIAKNFPLSKVSISSKVTYVGVVVAAPALGKGIQMOQLEMOON 420
 QY 424 VAQFOKEVGKLOAAADMTSMFTQFMOQASKIASKOTGESNEMTORATKLGIAQILKAYAI 483
 DB 421 VAQFOKEVGKLOAAADMTSMFTQFMOQASKIASKOTGESNEMTORATKLGIAQILKAYAI 480
 QY 484 SGATAGA 490
 DB 481 SGATAGA 487

RESULT 2
 US-08-809-326A-15
 Sequence 15, Application US/08809326A
 Patent No. 6165478
 GENERAL INFORMATION:
 APPLICANT: IZUTSU, HIROSHI
 APPLICANT: OBARA, KAZUHIKO

Tue Jan 28

THIS PAGE BLANK (UBPTD)

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 16:35:23 ; Search time 15 Seconds
(without alignments)
972.918 Million cell updates/sec

Title: US-09-889-314-2

Perfect score: 496
Sequence: 1 D1NMSSSSSGPDNOKNIMS.....LKAYAIISGAIAGAKHTNMF 496

Scoring table:

Gapop 60.0 , Gapext 60.0

Searched: 262574 seqs, 29422922 residues

Word size : 0

Total number of hits satisfying chosen parameters: 262574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database :

- Issued Patents AA:*
- 1: /cgn2_6/ptodata/1/iaa/5a_COMB.pep.*
 - 2: /cgn2_6/ptodata/1/iaa/5a_COMB.pep.*
 - 3: /cgn2_6/ptodata/1/iaa/5a_COMB.pep.*
 - 4: /cgn2_6/ptodata/1/iaa/5a_COMB.pep.*
 - 5: /cgn2_6/ptodata/1/iaa/5a_COMB.pep.*
 - 6: /cgn2_6/ptodata/1/iaa/5a_COMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	487	98.2	488	US-08-809-326A-1	Sequence 1, Appl
2	487	98.2	488	US-08-809-326A-1	Sequence 1, Appl
3	487	98.2	488	US-08-809-326A-1	Sequence 1, Appl
4	259	52.2	259	US-08-809-326A-2	Sequence 2, Appl
5	259	52.2	259	US-08-809-326A-2	Sequence 2, Appl
6	259	52.2	259	US-08-809-326A-2	Sequence 2, Appl
7	8	1.6	27	US-09-135-994-11	Sequence 11, Appl
8	8	1.6	28	US-08-425-069-50	Sequence 50, Appl
9	8	1.6	28	US-08-317-844B-50	Sequence 50, Appl
10	8	1.6	30	US-08-425-069-61	Sequence 61, Appl
11	8	1.6	30	US-08-317-844B-61	Sequence 61, Appl
12	8	1.6	33	US-08-425-069-6	Sequence 6, Appl
13	8	1.6	33	US-08-317-844B-6	Sequence 6, Appl
14	8	1.6	111	US-08-185-414E-2	Sequence 2, Appl
15	8	1.6	129	US-09-135-994-12	Sequence 12, Appl
16	8	1.6	164	US-08-911-319A-1	Sequence 1, Appl
17	8	1.6	164	US-08-911-319A-1	Sequence 1, Appl
18	8	1.6	205	US-08-775-009-37	Sequence 37, Appl
19	8	1.6	331	US-08-364-081-3	Sequence 3, Appl
20	8	1.6	331	US-08-364-081-3	Sequence 3, Appl
21	8	1.6	331	US-08-364-081-3	Sequence 3, Appl
22	8	1.6	625	US-09-347-801-18	Sequence 18, Appl
23	8	1.6	651	US-08-556-978B-19	Sequence 19, Appl
24	8	1.6	651	US-08-556-978B-19	Sequence 19, Appl
25	8	1.6	718	US-08-425-069-2	Sequence 2, Appl
26	8	1.6	747	US-08-317-844B-2	Sequence 2, Appl
27	8	1.6	1400	US-08-630-915A-37	Sequence 37, Appl

28	8	1.6	2088	US-09-548-372D-13	Sequence 13, Appl
29	8	1.6	2088	US-09-548-372D-13	Sequence 13, Appl
30	8	1.6	2211	US-09-738-884-1	Sequence 1, Appl
31	8	1.6	2508	US-09-627-650B-7	Sequence 7, Appl
32	8	1.6	2508	US-09-627-650B-7	Sequence 7, Appl
33	8	1.6	2544	US-09-436-063C-3	Sequence 3, Appl
34	8	1.6	2544	US-09-436-063C-3	Sequence 3, Appl
35	8	1.6	2601	US-09-627-650B-9	Sequence 9, Appl
36	8	1.6	2601	US-09-627-650B-9	Sequence 9, Appl
37	8	1.6	3788	US-09-336-447A-76	Sequence 76, Appl
38	8	1.6	3788	US-09-336-447A-76	Sequence 76, Appl
39	8	1.4	21	US-08-425-069-5	Sequence 5, Appl
40	8	1.4	21	US-08-317-844B-5	Sequence 5, Appl
41	8	1.4	23	US-07-987-286-3	Sequence 3, Appl
42	8	1.4	23	US-08-614-626-3	Sequence 3, Appl
43	8	1.4	25	US-07-987-286-11	Sequence 11, Appl
44	8	1.4	25	US-07-987-286-24	Sequence 24, Appl
45	8	1.4	25	US-08-425-069-38	Sequence 38, Appl
46	8	1.4	25	US-08-614-626-11	Sequence 11, Appl
47	8	1.4	25	US-08-317-844B-38	Sequence 38, Appl
48	8	1.4	25	US-08-317-844B-38	Sequence 38, Appl
49	8	1.4	27	US-08-425-069-47	Sequence 47, Appl
50	8	1.4	27	US-08-425-069-51	Sequence 51, Appl
51	8	1.4	27	US-08-425-069-55	Sequence 55, Appl
52	8	1.4	27	US-08-317-844B-47	Sequence 47, Appl
53	8	1.4	27	US-08-317-844B-51	Sequence 51, Appl
54	8	1.4	27	US-08-317-844B-55	Sequence 55, Appl
55	8	1.4	27	US-08-556-978B-29	Sequence 29, Appl
56	8	1.4	27	US-08-556-978B-92	Sequence 92, Appl
57	8	1.4	28	US-08-425-069-40	Sequence 40, Appl
58	8	1.4	28	US-08-425-069-43	Sequence 43, Appl
59	8	1.4	28	US-08-317-844B-40	Sequence 40, Appl
60	8	1.4	28	US-08-317-844B-43	Sequence 43, Appl
61	8	1.4	30	US-08-425-069-41	Sequence 41, Appl
62	8	1.4	30	US-08-425-069-48	Sequence 48, Appl
63	8	1.4	30	US-08-425-069-53	Sequence 53, Appl
64	8	1.4	30	US-08-425-069-58	Sequence 58, Appl
65	8	1.4	30	US-08-425-069-60	Sequence 60, Appl
66	8	1.4	30	US-08-317-844B-41	Sequence 41, Appl
67	8	1.4	30	US-08-317-844B-48	Sequence 48, Appl
68	8	1.4	30	US-08-317-844B-53	Sequence 53, Appl
69	8	1.4	30	US-08-317-844B-58	Sequence 58, Appl
70	8	1.4	30	US-08-317-844B-60	Sequence 60, Appl
71	8	1.4	30	US-08-556-978B-32	Sequence 32, Appl
72	8	1.4	30	US-08-556-978B-75	Sequence 75, Appl
73	8	1.4	30	US-08-556-978B-95	Sequence 95, Appl
74	8	1.4	31	US-08-425-069-46	Sequence 46, Appl
75	8	1.4	31	US-08-317-844B-46	Sequence 46, Appl
76	8	1.4	31	US-08-556-978B-26	Sequence 26, Appl
77	8	1.4	31	US-08-556-978B-69	Sequence 69, Appl
78	8	1.4	31	US-08-556-978B-89	Sequence 89, Appl
79	8	1.4	32	US-08-425-069-44	Sequence 44, Appl
80	8	1.4	32	US-08-317-844B-44	Sequence 44, Appl
81	8	1.4	33	US-08-425-069-7	Sequence 7, Appl
82	8	1.4	33	US-08-317-844B-7	Sequence 7, Appl
83	8	1.4	34	US-08-425-069-42	Sequence 42, Appl
84	8	1.4	34	US-08-425-069-52	Sequence 52, Appl
85	8	1.4	34	US-08-425-069-54	Sequence 54, Appl
86	8	1.4	34	US-08-317-844B-42	Sequence 42, Appl
87	8	1.4	34	US-08-317-844B-52	Sequence 52, Appl
88	8	1.4	34	US-08-317-844B-54	Sequence 54, Appl
89	8	1.4	34	US-08-317-844B-54	Sequence 54, Appl
90	8	1.4	58	US-08-279-058B-23	Sequence 23, Appl
91	8	1.4	58	US-08-828-323-23	Sequence 23, Appl
92	8	1.4	76	US-08-556-978B-22	Sequence 22, Appl
93	8	1.4	85	US-08-556-978B-20	Sequence 20, Appl
94	8	1.4	101	US-08-556-978B-13	Sequence 13, Appl
95	8	1.4	101	US-08-556-978B-62	Sequence 62, Appl
96	8	1.4	101	US-09-247-806-3	Sequence 3, Appl
97	8	1.4	101	US-09-247-806-5	Sequence 5, Appl
98	8	1.4	101	US-09-247-806-7	Sequence 7, Appl
99	8	1.4	102	US-08-273-146-65	Sequence 65, Appl
100	8	1.4	102	US-08-974-022-53	Sequence 53, Appl

101	7	1.4	102	4	US-08-795-445A-53	Sequence 53, Appl	174	7	1.4	445	3	US-09-165-543-5	Sequence 5, Appl
102	7	1.4	102	4	US-08-795-447A-53	Sequence 53, Appl	175	7	1.4	445	3	US-09-167-354-7	Sequence 7, Appl
103	7	1.4	102	4	US-08-974-186-53	Sequence 53, Appl	176	7	1.4	445	4	US-09-642-855-7	Sequence 7, Appl
104	7	1.4	102	4	US-08-795-446B-53	Sequence 2, Appl	177	7	1.4	445	4	US-09-642-514-7	Sequence 6, Appl
105	7	1.4	109	3	US-08-527-044-2	Sequence 2, Appl	178	7	1.4	457	2	US-08-882-704A-6	Sequence 2, Appl
106	7	1.4	109	3	US-09-013-780-2	Sequence 39, Appl	179	7	1.4	457	2	US-09-151-957-6	Sequence 6, Appl
107	7	1.4	111	4	US-09-091-725-39	Sequence 85, Appl	180	7	1.4	483	3	US-09-263-023-2	Sequence 2, Appl
108	7	1.4	113	2	US-08-256-568B-85	Sequence 8, Appl	181	7	1.4	483	3	US-09-471-867-2	Sequence 2, Appl
109	7	1.4	113	4	US-09-038-369B-85	Sequence 20, Appl	182	7	1.4	544	4	US-09-532-594B-18	Sequence 17, Appl
110	7	1.4	127	4	US-09-329-884-8	Sequence 18, Appl	183	7	1.4	548	2	US-08-871-866A-17	Sequence 23, Appl
111	7	1.4	132	1	US-08-611-107-20	Sequence 20, Appl	184	7	1.4	548	2	US-09-018-266A-17	Sequence 10, Appl
112	7	1.4	132	2	US-08-422-560A-20	Sequence 18, App	185	7	1.4	548	3	US-08-871-267B-23	Sequence 4, Appl
113	7	1.4	132	4	US-08-468-793-20	Sequence 18, App	186	7	1.4	548	3	US-09-618-419-23	Sequence 16, Appl
114	7	1.4	147	4	US-08-751-359-18	Sequence 108, App	187	7	1.4	550	4	US-09-387-574-10	Sequence 23, Appl
115	7	1.4	147	4	US-08-907-146-18	Sequence 18, App	188	7	1.4	572	3	US-09-668-096-10	Sequence 10, Appl
116	7	1.4	154	3	US-08-486-099-108	Sequence 108, App	189	7	1.4	572	3	US-09-040-681A-4	Sequence 4, Appl
117	7	1.4	154	3	US-08-360-107A-118	Sequence 108, App	190	7	1.4	572	4	US-09-497-897-4	Sequence 4, Appl
118	7	1.4	154	3	US-08-484-223B-108	Sequence 108, App	191	7	1.4	598	4	US-09-532-594B-16	Sequence 16, Appl
119	7	1.4	154	3	US-08-919-597-108	Sequence 108, App	192	7	1.4	604	4	US-08-536-978B-63	Sequence 63, Appl
120	7	1.4	154	3	US-08-475-668A-108	Sequence 108, App	193	7	1.4	606	4	US-08-556-978B-21	Sequence 21, Appl
121	7	1.4	154	3	US-08-485-551A-108	Sequence 108, App	194	7	1.4	606	4	US-08-556-978B-23	Sequence 23, Appl
122	7	1.4	154	3	US-08-471-913A-108	Sequence 108, App	195	7	1.4	606	4	US-09-247-806-6	Sequence 4, Appl
123	7	1.4	154	4	US-08-485-264A-108	Sequence 108, App	196	7	1.4	606	4	US-09-247-806-6	Sequence 4, Appl
124	7	1.4	154	4	US-08-474-349A-108	Sequence 108, App	197	7	1.4	606	4	US-08-365-981-15	Sequence 8, Appl
125	7	1.4	180	6	5273901-7	Patent No. 5273901	198	7	1.4	625	1	US-09-232-191-33	Sequence 15, Appl
126	7	1.4	180	6	5482709-6	Patent No. 5482709	199	7	1.4	643	4	US-09-232-200-83	Sequence 83, Appl
127	7	1.4	227	2	US-08-164-292B-27	Sequence 27, Appl	200	7	1.4	643	4	US-09-232-197-83	Sequence 83, Appl
128	7	1.4	227	2	US-08-845-623-27	Sequence 27, Appl	201	7	1.4	643	4	US-09-232-201-83	Sequence 83, Appl
129	7	1.4	227	3	US-08-815-927-27	Sequence 27, Appl	202	7	1.4	734	4	US-09-533-554B-4	Sequence 4, Appl
130	7	1.4	227	4	US-09-103-330-27	Sequence 27, Appl	203	7	1.4	734	4	US-09-533-554B-4	Sequence 4, Appl
131	7	1.4	227	4	US-09-435-242-27	Sequence 27, Appl	204	7	1.4	804	4	US-09-067-091-2	Sequence 2, Appl
132	7	1.4	235	4	US-09-134-001C-4827	Sequence 4827, Ap	205	7	1.4	805	4	US-09-134-001C-4821	Sequence 4821, Ap
133	7	1.4	241	1	US-07-987-286-17	Sequence 17, Appl	206	7	1.4	924	1	US-08-481-130A-28	Sequence 28, Appl
134	7	1.4	241	2	US-08-614-626-17	Sequence 17, Appl	207	7	1.4	924	1	US-08-656-984A-28	Sequence 28, Appl
135	7	1.4	243	1	US-07-690-983D-12	Sequence 12, Appl	208	7	1.4	924	1	US-08-485-604-28	Sequence 28, Appl
136	7	1.4	243	1	US-08-133-804-6	Sequence 6, Appl	209	7	1.4	924	2	US-08-485-595-28	Sequence 2, Appl
137	7	1.4	243	2	US-08-461-883-6	Sequence 6, Appl	210	7	1.4	924	2	US-08-485-595-28	Sequence 2, Appl
138	7	1.4	243	2	US-08-461-886-6	Sequence 6, Appl	211	7	1.4	992	1	US-08-137-499A-1	Sequence 1, Appl
139	7	1.4	247	2	US-08-356-786-4	Sequence 2, Appl	212	7	1.4	992	1	US-08-462-847-1	Sequence 1, Appl
140	7	1.4	247	2	US-08-951-822-23	Sequence 23, Appl	213	7	1.4	1127	4	US-09-150-460B-11	Sequence 11, Appl
141	7	1.4	247	2	US-08-705-245-4	Sequence 23, Appl	214	7	1.4	1276	4	US-08-937-236-3	Sequence 3, Appl
142	7	1.4	247	4	US-09-368-951-23	Sequence 23, Appl	215	7	1.4	1276	4	US-08-937-236-6	Sequence 6, Appl
143	7	1.4	252	1	US-08-462-169B-23	Sequence 23, Appl	216	7	1.4	1291	4	US-08-569-214-5	Sequence 5, Appl
144	7	1.4	252	3	US-09-103-079-2	Sequence 9, Appl	217	7	1.4	1292	4	US-08-569-214-6	Sequence 6, Appl
145	7	1.4	252	3	US-08-851-843A-174	Sequence 23, Appl	218	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
146	7	1.4	261	3	US-08-974-549A-293	Sequence 23, Appl	219	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
147	7	1.4	261	4	US-08-854-050-174	Sequence 174, App	220	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
148	7	1.4	261	4	US-09-430-323-114	Sequence 174, App	221	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
149	7	1.4	261	4	US-09-372-422A-32	Sequence 32, App	222	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
150	7	1.4	262	4	US-09-455-960-9	Sequence 9, Appl	223	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
151	7	1.4	282	4	US-08-146-421-5	Sequence 5, Appl	224	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
152	7	1.4	287	1	US-09-178-610-6	Sequence 6, Appl	225	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
153	7	1.4	305	3	US-09-540-014-9	Sequence 9, Appl	226	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
154	7	1.4	332	4	US-08-984-618-15	Sequence 24, Appl	227	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
155	7	1.4	351	4	US-09-347-975-2	Sequence 15, Appl	228	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
156	7	1.4	362	3	US-07-857-224B-80	Sequence 2, Appl	229	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
157	7	1.4	374	2	US-07-857-224B-81	Sequence 80, Appl	230	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
158	7	1.4	374	2	US-07-857-224B-82	Sequence 82, Appl	231	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
159	7	1.4	374	2	US-07-857-224B-83	Sequence 82, Appl	232	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
160	7	1.4	374	2	US-07-857-224B-84	Sequence 82, Appl	233	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
161	7	1.4	374	2	US-07-857-224B-85	Sequence 82, Appl	234	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
162	7	1.4	374	2	US-07-857-224B-86	Sequence 82, Appl	235	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
163	7	1.4	374	2	US-07-857-224B-87	Sequence 82, Appl	236	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
164	7	1.4	374	2	US-07-857-224B-88	Sequence 82, Appl	237	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
165	7	1.4	374	2	US-07-857-224B-89	Sequence 82, Appl	238	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
166	7	1.4	376	4	US-09-056-356-202	Sequence 202, App	239	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
167	7	1.4	376	4	US-09-072-596-197	Sequence 197, App	240	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
168	7	1.4	425	3	US-09-109-204-31	Sequence 31, Appl	241	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
169	7	1.4	444	1	US-07-881-075-3	Sequence 3, Appl	242	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
170	7	1.4	444	1	US-08-120-827-3	Sequence 3, Appl	243	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
171	7	1.4	444	1	US-08-478-675-3	Sequence 2, Appl	244	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
172	7	1.4	445	2	US-08-985-090-2	Sequence 2, Appl	245	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl
173	7	1.4	445	3	US-09-165-543-2	Sequence 2, Appl	246	7	1.4	1292	4	US-08-937-236-5	Sequence 5, Appl

247	6	1.2	12	2	US-08-630-645-23	Sequence 23, Appl	320	6	1.2	28	4	US-08-957-001B-25	Sequence 25, Appl
248	6	1.2	12	2	US-08-926-789-7	Sequence 7, Appl	321	6	1.2	28	4	US-09-496-301-25	Sequence 25, Appl
249	6	1.2	12	2	US-08-818-253-41	Sequence 41, Appl	322	6	1.2	28	4	US-08-556-978B-72	Sequence 72, Appl
250	6	1.2	12	3	US-09-166-750-1	Sequence 1, Appl	323	6	1.2	28	4	US-09-288-143-146	Sequence 146, Appl
251	6	1.2	12	3	US-09-166-093-1	Sequence 1, Appl	324	6	1.2	29	1	US-08-244-701B-51	Sequence 51, Appl
252	6	1.2	12	3	US-09-172-019-1	Sequence 1, Appl	325	6	1.2	29	4	US-09-128-440-9	Sequence 9, Appl
253	6	1.2	12	3	US-09-166-094-1	Sequence 1, Appl	326	6	1.2	29	4	US-09-823-494-9	Sequence 9, Appl
254	6	1.2	12	4	US-08-818-252-41	Sequence 41, Appl	327	6	1.2	29	4	US-09-076-721-51	Sequence 51, Appl
255	6	1.2	12	4	US-09-362-805-3	Sequence 3, Appl	328	6	1.2	30	3	US-08-256-747C-44	Sequence 44, Appl
256	6	1.2	12	4	US-09-173-190-3	Sequence 3, Appl	329	6	1.2	30	4	US-08-834-130A-44	Sequence 44, Appl
257	6	1.2	12	4	US-09-069-821-6	Sequence 6, Appl	330	6	1.2	32	1	US-07-828-450-41	Sequence 41, Appl
258	6	1.2	12	4	US-08-766-596A-23	Sequence 23, Appl	331	6	1.2	32	1	US-08-297-330-2	Sequence 2, Appl
259	6	1.2	12	4	US-09-316-919-57	Sequence 57, Appl	332	6	1.2	32	1	US-08-451-240-3	Sequence 2, Appl
260	6	1.2	12	5	PCT-US93-11138-7	Sequence 7, Appl	333	6	1.2	32	1	US-08-451-240-22	Sequence 22, Appl
261	6	1.2	12	5	PCT-US96-10220-23	Sequence 23, Appl	334	6	1.2	32	2	US-08-470-846A-3	Sequence 3, Appl
262	6	1.2	13	4	US-08-942-456-4	Sequence 4, Appl	335	6	1.2	32	3	US-08-954-915A-1	Sequence 1, Appl
263	6	1.2	13	4	US-08-525-539A-66	Sequence 66, Appl	336	6	1.2	32	3	US-08-954-915A-48	Sequence 48, Appl
264	6	1.2	13	4	US-09-510-721-4	Sequence 4, Appl	337	6	1.2	32	3	US-08-942-332B-1	Sequence 1, Appl
265	6	1.2	13	5	PCT-US95-04121-50	Sequence 50, Appl	338	6	1.2	32	4	US-08-942-456-1	Sequence 1, Appl
266	6	1.2	14	1	US-07-961-837-8	Sequence 8, Appl	339	6	1.2	32	4	US-09-510-721-1	Sequence 1, Appl
267	6	1.2	14	1	US-08-556-823-3	Sequence 3, Appl	340	6	1.2	32	4	US-08-924-6296-67	Sequence 67, Appl
268	6	1.2	14	1	US-08-556-823-4	Sequence 4, Appl	341	6	1.2	32	4	US-09-466-268B-6	Sequence 6, Appl
269	6	1.2	14	2	US-08-373-190-3	Sequence 3, Appl	342	6	1.2	32	5	PCT-US94-02391-2	Sequence 2, Appl
270	6	1.2	14	2	US-08-373-190-5	Sequence 5, Appl	343	6	1.2	32	5	PCT-US94-12591-3	Sequence 3, Appl
271	6	1.2	14	2	US-08-438-190A-3	Sequence 3, Appl	344	6	1.2	32	5	PCT-US94-12591-22	Sequence 22, Appl
272	6	1.2	14	2	US-08-438-190A-5	Sequence 5, Appl	345	6	1.2	33	1	US-08-209-747-19	Sequence 19, Appl
273	6	1.2	14	2	US-08-818-253-46	Sequence 46, Appl	346	6	1.2	33	1	US-08-556-823-7	Sequence 7, Appl
274	6	1.2	14	3	US-08-350-213-3	Sequence 3, Appl	347	6	1.2	33	1	US-08-458-298-19	Sequence 19, Appl
275	6	1.2	14	3	US-08-350-213-5	Sequence 5, Appl	348	6	1.2	33	1	US-08-244-701B-1	Sequence 1, Appl
276	6	1.2	14	3	US-09-287-145A-3	Sequence 3, Appl	349	6	1.2	33	1	US-08-244-701B-2	Sequence 2, Appl
277	6	1.2	14	3	US-09-287-145A-5	Sequence 5, Appl	350	6	1.2	33	4	US-09-128-450-8	Sequence 8, Appl
278	6	1.2	14	4	US-09-046-992-6	Sequence 6, Appl	351	6	1.2	33	4	US-09-128-450-17	Sequence 17, Appl
279	6	1.2	14	4	US-08-818-252-46	Sequence 46, Appl	352	6	1.2	33	4	US-09-128-450-23	Sequence 23, Appl
280	6	1.2	14	4	US-09-362-803-8	Sequence 8, Appl	353	6	1.2	33	4	US-09-128-450-29	Sequence 29, Appl
281	6	1.2	14	4	US-09-173-190-8	Sequence 8, Appl	354	6	1.2	33	4	US-09-823-494-8	Sequence 8, Appl
282	6	1.2	14	4	US-09-556-111-3	Sequence 3, Appl	355	6	1.2	33	4	US-09-823-494-17	Sequence 17, Appl
283	6	1.2	14	4	US-09-556-111-5	Sequence 5, Appl	356	6	1.2	33	4	US-08-823-494-23	Sequence 23, Appl
284	6	1.2	14	4	US-08-766-596A-34	Sequence 34, Appl	357	6	1.2	33	4	US-09-823-494-29	Sequence 29, Appl
285	6	1.2	14	4	US-09-316-919-62	Sequence 62, Appl	358	6	1.2	33	4	US-09-076-721-1	Sequence 1, Appl
286	6	1.2	15	2	US-08-528-057-18	Sequence 18, Appl	359	6	1.2	33	4	US-09-076-721-2	Sequence 2, Appl
287	6	1.2	15	2	US-08-373-190-4	Sequence 4, Appl	360	6	1.2	33	4	US-08-766-596A-37	Sequence 37, Appl
288	6	1.2	15	2	US-08-438-190A-4	Sequence 4, Appl	361	6	1.2	34	4	US-08-556-978B-1	Sequence 1, Appl
289	6	1.2	15	3	US-08-350-215-4	Sequence 4, Appl	362	6	1.2	36	1	US-08-209-747-31	Sequence 31, Appl
290	6	1.2	15	3	US-09-287-145A-4	Sequence 4, Appl	363	6	1.2	36	1	US-08-458-298-31	Sequence 31, Appl
291	6	1.2	15	4	US-09-556-111-4	Sequence 4, Appl	364	6	1.2	37	1	US-08-314-596-6	Sequence 6, Appl
292	6	1.2	16	2	US-08-308-494A-2	Sequence 2, Appl	365	6	1.2	37	1	US-08-314-596-22	Sequence 22, Appl
293	6	1.2	16	4	US-08-469-260A-236	Sequence 236, App	366	6	1.2	37	1	US-08-320-982-6	Sequence 6, Appl
294	6	1.2	18	1	US-08-240-712-33	Sequence 33, Appl	367	6	1.2	37	1	US-08-320-982-22	Sequence 22, Appl
295	6	1.2	18	1	US-08-443-890-33	Sequence 33, Appl	368	6	1.2	37	3	US-08-819-037-6	Sequence 6, Appl
296	6	1.2	18	3	US-09-053-617A-1	Sequence 1, Appl	369	6	1.2	37	3	US-08-819-037-22	Sequence 22, Appl
297	6	1.2	18	4	US-09-628-588-1	Sequence 1, Appl	370	6	1.2	37	4	US-08-900-230-54	Sequence 54, Appl
298	6	1.2	19	1	US-08-556-823-5	Sequence 5, Appl	371	6	1.2	37	4	US-09-045-940-6	Sequence 6, Appl
299	6	1.2	19	1	US-08-244-701B-3	Sequence 3, Appl	372	6	1.2	37	4	US-09-045-940-22	Sequence 22, Appl
300	6	1.2	19	1	US-08-244-701B-5	Sequence 5, Appl	373	6	1.2	38	1	US-08-318-193-3	Sequence 3, Appl
301	6	1.2	19	4	US-09-076-721-3	Sequence 3, Appl	374	6	1.2	38	6	5200327-14	Patent No. 5200327
302	6	1.2	19	4	US-09-076-721-5	Sequence 5, Appl	375	6	1.2	39	1	US-08-318-193-71	Sequence 71, Appl
303	6	1.2	21	1	US-08-164-151-15	Sequence 15, Appl	376	6	1.2	41	3	US-08-491-954-38	Sequence 38, Appl
304	6	1.2	21	3	US-08-942-332B-4	Sequence 4, Appl	377	6	1.2	41	3	US-08-993-165-9	Sequence 9, Appl
305	6	1.2	21	4	US-09-392-055-12	Sequence 12, Appl	378	6	1.2	41	4	US-09-540-448-9	Sequence 9, Appl
306	6	1.2	21	4	US-08-766-596A-35	Sequence 35, Appl	379	6	1.2	41	4	US-09-466-268B-1	Patent No. 5514550
307	6	1.2	22	4	US-08-918-148-39	Sequence 39, Appl	380	6	1.2	41	6	5514590-6	Patent No. 5514590
308	6	1.2	23	4	US-09-128-450-7	Sequence 7, Appl	381	6	1.2	42	1	US-08-377-687-39	Sequence 39, Appl
309	6	1.2	23	4	US-09-823-494-7	Sequence 7, Appl	382	6	1.2	42	2	US-08-777-192-39	Sequence 39, Appl
310	6	1.2	24	4	US-08-983-607-33	Sequence 33, Appl	383	6	1.2	42	4	US-08-971-982-39	Sequence 39, Appl
311	6	1.2	25	1	US-08-318-193-74	Sequence 74, Appl	384	6	1.2	43	1	US-08-078-683A-10	Sequence 10, Appl
312	6	1.2	25	4	US-09-466-268B-7	Sequence 7, Appl	385	6	1.2	43	1	US-08-078-683A-11	Sequence 11, Appl
313	6	1.2	26	2	US-08-850-910A-1	Sequence 1, Appl	386	6	1.2	43	1	US-08-178-477B-21	Sequence 21, Appl
314	6	1.2	27	1	US-08-425-069-56	Sequence 56, Appl	387	6	1.2	44	3	US-09-053-197A-55	Sequence 55, Appl
315	6	1.2	27	1	US-08-425-069-59	Sequence 59, Appl	388	6	1.2	44	4	US-09-085-761A-60	Sequence 60, Appl
316	6	1.2	27	2	US-08-317-844B-56	Sequence 56, Appl	389	6	1.2	45	1	US-08-318-193-12	Sequence 12, Appl
317	6	1.2	27	2	US-08-317-844B-59	Sequence 59, Appl	390	6	1.2	45	1	US-08-451-947-97	Sequence 97, Appl
318	6	1.2	28	1	US-08-276-213-4	Sequence 4, Appl	391	6	1.2	45	2	US-08-424-826A-97	Sequence 97, Appl
319	6	1.2	28	2	US-08-640-847C-38	Sequence 38, Appl	392	6	1.2	45	3	US-08-928-694-97	Sequence 97, Appl

393	6	1.2	45	4	US-08-900-230-7	Sequence 7, Appli	466	6	1.2	111	1	US-08-466-886-43	Sequence 43, Appli
394	6	1.2	45	5	PCT-US91-06950-97	Sequence 97, Appli	467	6	1.2	111	4	US-08-466-617-43	Sequence 43, Appli
395	6	1.2	47	1	US-08-209-747-24	Sequence 24, Appli	468	6	1.2	112	1	US-08-211-202-133	Sequence 133, App
396	6	1.2	47	1	US-08-209-747-26	Sequence 26, Appli	469	6	1.2	112	3	US-08-991-813-4	Sequence 4, Appli
397	6	1.2	47	1	US-08-458-298-24	Sequence 24, Appli	470	6	1.2	113	2	US-08-956-568B-66	Sequence 86, Appli
398	6	1.2	47	1	US-08-458-298-26	Sequence 26, Appli	471	6	1.2	113	2	US-08-956-568B-66	Sequence 86, Appli
399	6	1.2	48	1	US-08-209-747-14	Sequence 14, Appli	472	6	1.2	113	3	US-08-513-974B-38	Sequence 38, Appli
400	6	1.2	48	1	US-08-209-747-21	Sequence 21, Appli	473	6	1.2	115	3	US-08-615-192A-317	Sequence 317, App
401	6	1.2	48	1	US-08-458-298-11	Sequence 11, Appli	474	6	1.2	117	3	US-08-545-809A-107	Sequence 107, App
402	6	1.2	48	1	US-08-458-298-21	Sequence 21, Appli	475	6	1.2	117	3	US-08-301-922-12	Sequence 12, Appli
403	6	1.2	49	1	US-08-209-747-22	Sequence 22, Appli	476	6	1.2	118	3	US-08-545-809A-123	Sequence 123, App
404	6	1.2	49	1	US-08-458-298-22	Sequence 22, Appli	477	6	1.2	118	3	US-08-968-927-10	Sequence 10, Appli
405	6	1.2	50	1	US-08-318-193-24	Sequence 24, Appli	478	6	1.2	118	4	US-09-461-240-10	Sequence 10, Appli
406	6	1.2	50	1	US-08-209-747-27	Sequence 27, Appli	479	6	1.2	118	4	US-09-968-927-10	Sequence 10, Appli
407	6	1.2	50	1	US-08-458-298-27	Sequence 27, Appli	480	6	1.2	122	2	US-08-450-578-11	Sequence 11, Appli
408	6	1.2	50	4	US-08-905-223-488	Sequence 488, App	481	6	1.2	122	2	US-09-017-628-11	Sequence 11, Appli
409	6	1.2	50	4	US-09-262-656A-1	Sequence 1, Appli	482	6	1.2	122	2	US-09-014-860-11	Sequence 11, Appli
410	6	1.2	51	1	US-08-209-747-20	Sequence 20, Appli	483	6	1.2	122	2	US-08-450-363-11	Sequence 11, Appli
411	6	1.2	51	1	US-08-209-747-30	Sequence 30, Appli	484	6	1.2	122	2	US-08-078-683A-9	Sequence 9, Appli
412	6	1.2	51	1	US-08-458-298-20	Sequence 20, Appli	485	6	1.2	123	1	US-09-615-192A-321	Sequence 321, App
413	6	1.2	51	1	US-08-209-747-6	Sequence 6, Appli	486	6	1.2	125	4	US-09-160-246-16	Sequence 16, App
414	6	1.2	54	1	US-08-209-747-29	Sequence 29, Appli	487	6	1.2	127	4	US-09-291-922-4	Sequence 4, Appli
415	6	1.2	54	1	US-08-458-298-6	Sequence 6, Appli	488	6	1.2	131	4	US-08-513-974B-325	Sequence 325, App
416	6	1.2	54	1	US-08-209-747-28	Sequence 28, Appli	489	6	1.2	132	3	US-08-513-974B-326	Sequence 10, Appli
417	6	1.2	54	1	US-08-905-223-427	Sequence 29, Appli	490	6	1.2	132	3	US-09-452-239-10	Sequence 45, Appli
418	6	1.2	54	1	US-08-209-747-10	Sequence 10, Appli	491	6	1.2	133	4	US-08-850-910A-45	Sequence 48, Appli
419	6	1.2	55	1	US-08-209-747-32	Sequence 32, Appli	492	6	1.2	134	2	US-08-505-218-13	Sequence 13, Appli
420	6	1.2	55	1	US-08-458-298-12	Sequence 12, Appli	493	6	1.2	136	2	US-08-444-818-16	Sequence 16, Appli
421	6	1.2	55	1	US-08-209-747-32	Sequence 32, Appli	494	6	1.2	136	2	US-09-091-725-51	Sequence 51, Appli
422	6	1.2	55	1	US-08-458-298-25	Sequence 25, Appli	495	6	1.2	141	4	US-08-556-823-2	Sequence 2, Appli
423	6	1.2	55	1	US-08-458-298-32	Sequence 32, Appli	496	6	1.2	142	1	US-08-556-823-10	Sequence 10, Appli
424	6	1.2	55	1	US-08-209-747-28	Sequence 28, Appli	497	6	1.2	142	1	US-08-505-218-2	Sequence 2, Appli
425	6	1.2	56	1	US-08-556-823-1	Sequence 1, Appli	498	6	1.2	142	2	US-08-318-193-16	Sequence 16, Appli
426	6	1.2	56	1	US-08-556-823-9	Sequence 9, Appli	499	6	1.2	144	2	US-09-372-422A-46	Sequence 46, Appli
427	6	1.2	56	1	US-08-458-298-28	Sequence 28, Appli	500	6	1.2	150	4	US-08-318-947A-9	Sequence 9, Appli
428	6	1.2	56	1	PCT-US96-01720-10	Sequence 10, Appli	501	6	1.2	151	1	US-08-795-303-9	Sequence 33, Appli
429	6	1.2	71	5	PCT-US96-01720-11	Sequence 11, Appli	502	6	1.2	151	2	US-08-933-750C-33	Sequence 33, Appli
430	6	1.2	71	5	US-09-461-697-42	Sequence 42, Appli	503	6	1.2	151	2	US-09-234-613-33	Sequence 4, Appli
431	6	1.2	72	4	US-08-858-207A-368	Sequence 368, App	504	6	1.2	152	2	US-08-209-747-15	Sequence 15, Appli
432	6	1.2	80	1	US-08-209-747-12	Sequence 12, Appli	505	6	1.2	153	4	US-08-458-298-15	Sequence 15, Appli
433	6	1.2	80	1	US-08-528-057-8	Sequence 8, Appli	506	6	1.2	153	4	US-09-096-244-4	Sequence 4, Appli
434	6	1.2	81	2	US-09-014-969-32	Sequence 32, Appli	507	6	1.2	155	1	US-08-209-747-15	Sequence 15, Appli
435	6	1.2	81	2	US-08-528-057-5	Sequence 5, Appli	508	6	1.2	155	1	US-08-458-298-15	Sequence 15, Appli
436	6	1.2	81	4	US-09-134-001C-340	Sequence 340, Ap	509	6	1.2	155	4	US-09-370-838-210	Sequence 210, App
437	6	1.2	85	2	US-08-528-057-7	Sequence 7, Appli	510	6	1.2	156	4	US-08-933-855-196	Sequence 196, App
438	6	1.2	88	2	US-08-821-009-1	Sequence 1, Appli	511	6	1.2	159	3	US-09-093-855-7	Sequence 7, Appli
439	6	1.2	90	4	US-09-247-155-86	Sequence 86, Appli	512	6	1.2	159	4	US-09-540-448-7	Sequence 7, Appli
440	6	1.2	90	4	US-08-591-498-3	Sequence 98, Appli	513	6	1.2	159	4	US-09-615-192A-318	Sequence 318, App
441	6	1.2	91	1	US-08-528-057-15	Sequence 15, Appli	514	6	1.2	161	4	US-08-318-193-8	Sequence 8, Appli
442	6	1.2	92	2	US-08-821-009-1	Sequence 1, Appli	515	6	1.2	165	4	US-09-068-195-14	Sequence 14, Appli
443	6	1.2	92	2	US-08-528-057-7	Sequence 7, Appli	516	6	1.2	166	4	US-08-818-111-95	Sequence 11, Appli
444	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	517	6	1.2	166	4	US-08-818-111-100	Sequence 100, App
445	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	518	6	1.2	168	4	US-09-072-596-95	Sequence 95, Appli
446	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	519	6	1.2	168	4	US-08-310-416A-18	Sequence 18, Appli
447	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	520	6	1.2	169	4	US-08-888-171-18	Sequence 171, Appli
448	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	521	6	1.2	169	4	US-08-888-171-18	Sequence 18, Appli
449	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	522	6	1.2	169	4	US-08-888-171-18	Sequence 18, Appli
450	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	523	6	1.2	169	4	US-08-888-171-18	Sequence 18, Appli
451	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	524	6	1.2	169	4	US-08-888-171-18	Sequence 18, Appli
452	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	525	6	1.2	169	4	US-08-888-171-18	Sequence 18, Appli
453	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	526	6	1.2	169	4	US-08-888-171-18	Sequence 18, Appli
454	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	527	6	1.2	169	4	US-08-888-171-18	Sequence 18, Appli
455	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	528	6	1.2	169	4	US-08-888-171-18	Sequence 18, Appli
456	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	529	6	1.2	169	4	US-08-888-171-18	Sequence 18, Appli
457	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	530	6	1.2	169	4	US-08-888-171-18	Sequence 18, Appli
458	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	531	6	1.2	169	4	US-08-888-171-18	Sequence 18, Appli
459	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	532	6	1.2	169	4	US-08-888-171-18	Sequence 18, Appli
460	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	533	6	1.2	169	4	US-08-888-171-18	Sequence 18, Appli
461	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	534	6	1.2	169	4	US-08-888-171-18	Sequence 18, Appli
462	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	535	6	1.2	169	4	US-08-888-171-18	Sequence 18, Appli
463	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	536	6	1.2	169	4	US-08-888-171-18	Sequence 18, Appli
464	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	537	6	1.2	169	4	US-08-888-171-18	Sequence 18, Appli
465	6	1.2	92	2	US-08-461-697-38	Sequence 38, Appli	538	6	1.2	169	4	US-08-888-171-18	Sequence 18, Appli

539	6	1.2	187	4	US-09-056-556-63	Sequence 63, Appl	612	6	1.2	246	1	US-08-294-386C-3	Sequence 3, Appl1
540	6	1.2	187	4	US-09-072-596-64	Sequence 64, Appl	613	6	1.2	246	1	US-08-299-249A-1	Sequence 1, Appl1
541	6	1.2	190	2	US-08-466-033-38	Sequence 38, Appl	614	6	1.2	246	1	US-08-299-249A-10	Sequence 10, Appl
542	6	1.2	190	2	US-08-444-733-38	Sequence 38, Appl	615	6	1.2	246	1	US-08-590-708-1	Sequence 1, Appl1
543	6	1.2	190	2	US-08-464-134-38	Sequence 38, Appl	616	6	1.2	246	5	PCT-US94-06543-1	Sequence 1, Appl1
544	6	1.2	190	2	US-08-461-361-38	Sequence 38, Appl	617	6	1.2	246	5	PCT-US95-10224-1	Sequence 1, Appl1
545	6	1.2	190	2	US-08-485-910-38	Sequence 38, Appl	618	6	1.2	246	5	PCT-US95-10224-3	Sequence 3, Appl1
546	6	1.2	190	4	US-09-125-619-27	Sequence 27, Appl	619	6	1.2	247	1	US-08-324-977-44	Sequence 44, Appl
547	6	1.2	190	5	PCT-US95-06266-29	Sequence 29, Appl	620	6	1.2	247	2	US-08-384-616-44	Sequence 44, Appl
548	6	1.2	194	3	US-08-480-640A-196	Sequence 196, App	621	6	1.2	247	2	US-08-304-686A-44	Sequence 44, Appl
549	6	1.2	194	4	US-08-686-968C-196	Sequence 196, App	622	6	1.2	247	4	US-09-315-850-44	Sequence 44, Appl
550	6	1.2	194	4	US-08-488-237A-196	Sequence 196, App	623	6	1.2	250	4	US-09-134-001C-2869	Sequence 2869, Ap
551	6	1.2	194	4	US-08-375-992A-196	Sequence 196, App	624	6	1.2	251	1	US-08-209-747-8	Sequence 8, Appl1
552	6	1.2	195	4	US-09-288-143-144	Sequence 144, App	625	6	1.2	251	1	US-08-458-298-8	Sequence 8, Appl1
553	6	1.2	195	4	US-09-125-619-17	Sequence 17, Appl	626	6	1.2	253	1	US-08-242-188-2	Sequence 2, Appl1
554	6	1.2	197	4	US-09-125-619-25	Sequence 25, Appl	627	6	1.2	253	1	US-08-509-261A-2	Sequence 2, Appl1
555	6	1.2	198	1	US-08-278-091-16	Sequence 16, Appl	628	6	1.2	253	1	US-08-660-626-8	Sequence 8, Appl1
556	6	1.2	198	1	US-08-483-859-16	Sequence 16, Appl	629	6	1.2	253	1	US-08-692-892-2	Sequence 2, Appl1
557	6	1.2	198	1	US-08-472-173-16	Sequence 16, Appl	630	6	1.2	253	2	US-08-713-939A-2	Sequence 2, Appl1
558	6	1.2	198	2	US-08-487-167-16	Sequence 16, Appl	631	6	1.2	253	2	US-08-668-162A-22	Sequence 22, Appl
559	6	1.2	198	2	US-08-482-816-16	Sequence 16, Appl	632	6	1.2	253	4	US-09-031-168-8	Sequence 8, Appl1
560	6	1.2	198	2	US-08-296-149-16	Sequence 16, Appl	633	6	1.2	253	4	US-09-128-450-20	Sequence 20, Appl
561	6	1.2	198	2	US-08-801-499-16	Sequence 16, Appl	634	6	1.2	253	4	US-09-036-579-2	Sequence 2, Appl1
562	6	1.2	198	2	US-08-615-271-16	Sequence 16, Appl	635	6	1.2	253	4	US-09-823-494-20	Sequence 20, Appl
563	6	1.2	198	3	US-09-074-660-16	Sequence 16, Appl	636	6	1.2	253	4	US-09-550-374-2	Sequence 2, Appl1
564	6	1.2	198	3	US-09-106-468-16	Sequence 16, Appl	637	6	1.2	254	1	US-08-242-188-1	Sequence 1, Appl1
565	6	1.2	198	3	US-09-106-468-16	Sequence 16, Appl	638	6	1.2	254	1	US-08-236-918A-4	Sequence 4, Appl1
566	6	1.2	198	4	US-09-106-466A-16	Sequence 16, Appl	639	6	1.2	254	1	US-08-509-261A-1	Sequence 1, Appl1
567	6	1.2	198	4	US-09-106-467-16	Sequence 16, Appl	640	6	1.2	254	1	US-08-660-626-7	Sequence 7, Appl1
568	6	1.2	200	4	US-08-965-056-104	Sequence 104, App	641	6	1.2	254	1	US-08-692-892-1	Sequence 1, Appl1
569	6	1.2	200	4	US-09-562-737-107	Sequence 107, App	642	6	1.2	254	2	US-08-713-939A-1	Sequence 1, Appl1
570	6	1.2	201	4	US-09-134-001C-4319	Sequence 4319, Ap	643	6	1.2	254	2	US-08-668-162A-21	Sequence 21, Appl
571	6	1.2	202	4	US-09-199-637A-149	Sequence 149, App	644	6	1.2	254	4	US-09-031-168-7	Sequence 7, Appl1
572	6	1.2	208	4	US-09-128-450-18	Sequence 18, Appl	645	6	1.2	254	4	US-09-128-450-16	Sequence 16, Appl
573	6	1.2	208	4	US-09-823-494-18	Sequence 18, Appl	646	6	1.2	254	4	US-09-128-450-20	Sequence 20, Appl
574	6	1.2	209	4	US-09-134-001C-3637	Sequence 3637, Ap	647	6	1.2	254	4	US-09-128-450-28	Sequence 28, Appl
575	6	1.2	213	3	US-08-971-158-4	Sequence 4, Appl1	648	6	1.2	254	4	US-09-036-579-1	Sequence 1, Appl1
576	6	1.2	213	4	US-09-416-488-4	Sequence 4, Appl1	649	6	1.2	254	4	US-09-823-494-19	Sequence 19, Appl
577	6	1.2	214	2	US-08-984-172-1	Sequence 30, Appl1	650	6	1.2	254	4	US-09-823-494-26	Sequence 26, Appl
578	6	1.2	214	4	US-08-861-774E-30	Sequence 30, Appl1	651	6	1.2	254	4	US-09-823-494-28	Sequence 28, Appl
579	6	1.2	215	4	US-09-134-001C-4736	Sequence 4736, Ap	652	6	1.2	254	4	US-09-150-864A-4	Sequence 4, Appl1
580	6	1.2	216	4	US-09-125-619-46	Sequence 46, Appl	653	6	1.2	254	4	US-09-550-374-1	Sequence 1, Appl1
581	6	1.2	217	4	US-09-416-509C-1	Sequence 1, Appl1	654	6	1.2	255	1	US-08-242-188-4	Sequence 4, Appl1
582	6	1.2	219	1	US-08-463-115-91	Sequence 91, Appl	655	6	1.2	255	1	US-08-509-261A-4	Sequence 4, Appl1
583	6	1.2	219	1	US-08-465-388-91	Sequence 91, Appl	656	6	1.2	255	1	US-08-660-626-10	Sequence 10, Appl
584	6	1.2	222	3	US-08-924-747-22	Sequence 22, Appl	657	6	1.2	255	1	US-08-692-892-4	Sequence 4, Appl1
585	6	1.2	222	4	US-09-247-373B-22	Sequence 22, Appl	658	6	1.2	255	2	US-08-713-939A-4	Sequence 4, Appl1
586	6	1.2	222	4	US-09-296-715-22	Sequence 22, Appl	659	6	1.2	255	2	US-08-868-162A-24	Sequence 24, Appl
587	6	1.2	223	5	PCT-US94-14277-6	Sequence 6, Appl1	660	6	1.2	255	4	US-09-031-168-10	Sequence 10, Appl
588	6	1.2	225	4	US-09-134-001C-3215	Sequence 3215, Ap	661	6	1.2	255	4	US-09-036-579-4	Sequence 4, Appl1
589	6	1.2	230	1	US-08-118-469A-3	Sequence 3, Appl1	662	6	1.2	255	4	US-09-550-374-4	Sequence 4, Appl1
590	6	1.2	230	1	US-08-909-119-3	Sequence 3, Appl1	663	6	1.2	256	4	US-09-128-450-22	Sequence 22, Appl
591	6	1.2	230	2	US-08-928-692-30	Sequence 30, Appl	664	6	1.2	256	4	US-09-823-494-22	Sequence 22, Appl
592	6	1.2	230	4	US-08-768-373-4	Sequence 4, Appl1	665	6	1.2	258	4	US-08-957-130-13	Sequence 13, Appl
593	6	1.2	230	4	US-08-818-112-65	Sequence 65, Appl	666	6	1.2	260	4	US-09-452-239-12	Sequence 12, Appl
594	6	1.2	230	4	US-09-339-972-30	Sequence 30, Appl	667	6	1.2	262	4	US-08-154-364-41	Sequence 41, Appl
595	6	1.2	230	4	US-08-818-111-66	Sequence 66, Appl	668	6	1.2	263	1	US-07-721-761A-51	Sequence 51, Appl
596	6	1.2	230	4	US-09-056-556-65	Sequence 65, Appl	669	6	1.2	263	1	US-08-242-188-3	Sequence 3, Appl1
597	6	1.2	230	4	US-09-072-596-66	Sequence 66, Appl	670	6	1.2	263	1	US-08-509-261A-3	Sequence 3, Appl1
598	6	1.2	233	2	US-08-928-692-31	Sequence 31, Appl	671	6	1.2	263	1	US-08-660-626-9	Sequence 9, Appl1
599	6	1.2	233	4	US-09-339-972-31	Sequence 31, Appl	672	6	1.2	263	1	US-08-692-892-3	Sequence 3, Appl1
600	6	1.2	237	3	US-09-216-001-1	Sequence 1, Appl1	673	6	1.2	263	2	US-08-713-939A-3	Sequence 3, Appl1
601	6	1.2	237	4	US-08-878-862-1	Sequence 1, Appl1	674	6	1.2	263	2	US-08-868-162A-23	Sequence 23, Appl
602	6	1.2	239	6	5455030-13	Patent No. 5455030	675	6	1.2	263	4	US-09-031-168-9	Sequence 9, Appl1
603	6	1.2	240	3	US-08-942-001-9	Sequence 9, Appl1	676	6	1.2	263	4	US-09-036-579-3	Sequence 3, Appl1
604	6	1.2	240	4	US-09-337-386-9	Sequence 9, Appl1	677	6	1.2	263	4	US-09-550-374-3	Sequence 3, Appl1
605	6	1.2	240	4	US-09-846-922-9	Sequence 9, Appl1	678	6	1.2	264	4	US-09-128-450-21	Sequence 21, Appl
606	6	1.2	241	2	US-08-700-607-3	Sequence 3, Appl1	679	6	1.2	264	4	US-09-823-494-21	Sequence 21, Appl
607	6	1.2	242	6	5455030-17	Patent No. 5455030	680	6	1.2	265	1	US-08-468-163-13	Sequence 13, Appl
608	6	1.2	243	3	US-09-216-295-13	Sequence 13, Appl	681	6	1.2	265	2	US-08-393-996A-19	Sequence 19, Appl
609	6	1.2	245	4	US-08-918-148-78	Sequence 78, Appl	682	6	1.2	266	4	US-09-615-192A-394	Sequence 394, App
610	6	1.2	246	1	US-08-286-888B-1	Sequence 1, Appl1	683	6	1.2	266	6	5223391-7	Patent No. 5223391
611	6	1.2	246	1	US-08-294-386C-1	Sequence 1, Appl1	684	6	1.2	267	1	US-07-706-872-3	Sequence 3, Appl1

685	6	1.2	268	3	US-08-871-483-11	Sequence 11, Appl	758	6	1.2	342	4	US-09-129-033-2	Sequence 2, Appl
686	6	1.2	269	3	US-09-100-557-1	Sequence 1, Appl	759	6	1.2	344	1	US-07-941-523-24	Sequence 24, Appl
687	6	1.2	269	3	US-08-926-842B-59	Sequence 159, Appl	760	6	1.2	344	2	US-07-857-224B-93	Sequence 93, Appl
688	6	1.2	270	4	US-09-134-001C-3187	Sequence 3187, Ap	761	6	1.2	344	2	US-08-602-725-34	Sequence 34, Appl
689	6	1.2	273	4	US-09-008-882-2	Sequence 2, Appl	762	6	1.2	344	4	US-09-134-001C-3158	Sequence 3158, Ap
690	6	1.2	277	4	US-09-134-001C-5261	Sequence 5261, Ap	763	6	1.2	344	4	US-09-134-001C-4304	Sequence 4304, Ap
691	6	1.2	278	1	US-07-941-651-2	Sequence 2, Appl	764	6	1.2	345	1	US-08-171-382-6	Sequence 6
692	6	1.2	278	1	US-07-988-260B-3	Sequence 3, Appl	765	6	1.2	345	1	US-08-309-420-6	Sequence 6, Appl
693	6	1.2	278	1	US-08-188-582-9	Sequence 9, Appl	766	6	1.2	345	1	US-08-309-420-6	Sequence 6, Appl
694	6	1.2	278	1	US-08-279-996-2	Sequence 9, Appl	767	6	1.2	345	1	US-08-309-420-6	Sequence 6, Appl
695	6	1.2	278	1	US-08-646-715-6	Sequence 9, Appl	768	6	1.2	345	5	US-09-294-531B-31	Sequence 31, Appl
696	6	1.2	278	1	US-08-646-715-6	Sequence 9, Appl	769	6	1.2	345	5	PCT-US95-118978-6	Sequence 31, Appl
697	6	1.2	278	5	PCT-US96-03916-60	Sequence 60, Appl	770	6	1.2	347	4	US-09-144-914-6	Sequence 6, Appl
698	6	1.2	280	1	US-08-434-255-8	Sequence 72, Appl	771	6	1.2	347	4	US-09-144-914-6	Sequence 6, Appl
699	6	1.2	280	1	US-08-459-967-8	Sequence 8, Appl	772	6	1.2	348	4	US-09-216-295-16	Sequence 6, Appl
700	6	1.2	280	1	US-08-460-327-8	Sequence 8, Appl	773	6	1.2	349	4	US-09-161-241-8	Sequence 6, Appl
701	6	1.2	280	1	US-08-459-871-8	Sequence 8, Appl	774	6	1.2	351	1	US-08-470-202-46	Sequence 8, Appl
702	6	1.2	280	4	US-09-160-246-14	Sequence 14, Appl	775	6	1.2	351	2	US-08-468-055-46	Sequence 46, Appl
703	6	1.2	280	4	US-09-024-532-2	Sequence 2, Appl	776	6	1.2	351	2	US-08-468-055-46	Sequence 46, Appl
704	6	1.2	280	4	US-09-104-623A-2	Sequence 2, Appl	777	6	1.2	351	4	US-09-109-910-46	Sequence 46, Appl
705	6	1.2	280	4	US-09-019-532-2	Sequence 2, Appl	778	6	1.2	353	4	US-09-134-001C-4640	Sequence 46, Appl
706	6	1.2	280	4	US-09-417-359A-2	Sequence 2, Appl	779	6	1.2	353	4	US-08-153-848-28	Sequence 28, Appl
707	6	1.2	283	2	US-08-852-401-4	Sequence 4, Appl	780	6	1.2	355	1	US-08-153-848-32	Sequence 32, Appl
708	6	1.2	285	2	US-08-712-072C-5	Sequence 4, Appl	781	6	1.2	355	3	US-09-299-843A-32	Sequence 28, Appl
709	6	1.2	283	1	US-08-310-416A-16	Sequence 16, Appl	782	6	1.2	355	3	US-09-299-843A-32	Sequence 28, Appl
710	6	1.2	293	2	US-08-888-171-16	Sequence 16, Appl	783	6	1.2	355	4	US-09-088-337B-28	Sequence 28, Appl
711	6	1.2	296	1	US-07-783-705A-5	Sequence 5, Appl	784	6	1.2	355	4	US-09-088-337B-28	Sequence 28, Appl
712	6	1.2	298	3	US-08-961-871-10	Sequence 10, Appl	785	6	1.2	355	4	US-09-088-337B-28	Sequence 28, Appl
713	6	1.2	299	6	5514590-4	Patent No. 5514590	786	6	1.2	355	5	PCT-US93-11153-38	Sequence 32, Appl
714	6	1.2	308	2	US-09-047-026A-23	Sequence 23, Appl	787	6	1.2	356	2	US-08-461-990B-2	Sequence 2, Appl
715	6	1.2	309	2	US-08-405-175A-7	Sequence 7, Appl	788	6	1.2	357	4	US-08-638-911A-37	Sequence 37, Appl
716	6	1.2	309	2	US-08-405-175A-7	Sequence 7, Appl	789	6	1.2	357	4	US-08-638-911A-37	Sequence 37, Appl
717	6	1.2	309	3	US-08-508-761B-32	Sequence 32, Appl	790	6	1.2	358	3	US-08-586-165-2	Sequence 2, Appl
718	6	1.2	310	3	US-08-897-097-4	Sequence 4, Appl	791	6	1.2	360	3	US-08-899-437-7	Sequence 7, Appl
719	6	1.2	311	2	US-08-897-097-4	Sequence 32, Appl	792	6	1.2	360	4	US-09-126-121-7	Sequence 7, Appl
720	6	1.2	311	3	US-08-775-009-32	Sequence 3, Appl	793	6	1.2	362	3	US-08-850-338-4	Sequence 4, Appl
721	6	1.2	313	4	US-09-336-643A-81	Sequence 81, Appl	794	6	1.2	362	3	US-08-899-437-3	Sequence 3, Appl
722	6	1.2	315	4	US-09-166-205B-27	Sequence 27, Appl	795	6	1.2	362	4	US-09-126-121-3	Sequence 3, Appl
723	6	1.2	317	2	US-08-864-799-4	Sequence 4, Appl	796	6	1.2	363	4	US-09-301-665-4	Sequence 4, Appl
724	6	1.2	317	2	US-08-864-799-4	Sequence 4, Appl	797	6	1.2	370	1	US-08-434-225-6	Sequence 6, Appl
725	6	1.2	317	4	US-08-864-799-5	Sequence 4, Appl	798	6	1.2	370	1	US-08-434-225-6	Sequence 6, Appl
726	6	1.2	317	4	US-08-937-067-6	Sequence 6, Appl	799	6	1.2	370	1	US-08-459-967-6	Sequence 6, Appl
727	6	1.2	318	4	US-08-278-042C-3	Sequence 3, Appl	800	6	1.2	370	1	US-08-460-337-6	Sequence 6, Appl
728	6	1.2	318	4	US-09-060-756-727	Sequence 727, App	801	6	1.2	370	2	US-08-459-871-6	Sequence 6, Appl
729	6	1.2	321	6	US-08-937-972-3	Sequence 3, Appl	802	6	1.2	370	2	US-08-528-057-12	Sequence 12, Appl
730	6	1.2	321	6	5169835-17	Patent No. 5169835	803	6	1.2	370	4	US-09-199-737-2	Sequence 42, Appl
731	6	1.2	323	3	US-08-853-839-2	Sequence 2, Appl	804	6	1.2	370	4	US-08-900-230-2	Sequence 2, Appl
732	6	1.2	323	3	US-09-041-889-28	Sequence 28, Appl	805	6	1.2	370	4	US-09-058-333A-2	Sequence 2, Appl
733	6	1.2	323	3	US-09-615-192A-395	Sequence 395, App	806	6	1.2	373	2	US-09-724-224-2	Sequence 2, Appl
734	6	1.2	325	2	US-08-107-676-30	Sequence 30, Appl	807	6	1.2	373	2	US-08-528-057-44	Sequence 44, Appl
735	6	1.2	325	3	US-08-303-861-12	Sequence 12, Appl	808	6	1.2	374	1	US-08-625-876-2	Sequence 2, Appl
736	6	1.2	325	4	US-08-975-762-8	Sequence 8, Appl	809	6	1.2	374	1	US-08-625-876-4	Sequence 4, Appl
737	6	1.2	325	4	US-08-821-324-8	Sequence 8, Appl	810	6	1.2	374	1	US-08-625-876-6	Sequence 6, Appl
738	6	1.2	325	4	US-09-295-028-8	Sequence 8, Appl	811	6	1.2	375	2	US-07-857-224B-86	Sequence 86, Appl
739	6	1.2	325	4	US-09-106-582-8	Sequence 8, Appl	812	6	1.2	375	2	US-07-857-224B-86	Sequence 86, Appl
740	6	1.2	327	3	US-08-513-974B-372	Sequence 372, App	813	6	1.2	376	4	US-09-444-412-2	Sequence 2, Appl
741	6	1.2	328	3	US-08-459-046-2	Sequence 3, Appl	814	6	1.2	376	4	US-07-857-224B-88	Sequence 88, Appl
742	6	1.2	328	3	US-08-513-974B-39	Sequence 39, Appl	815	6	1.2	377	3	US-08-528-057-2	Sequence 2, Appl
743	6	1.2	328	3	US-08-513-974B-56	Sequence 56, Appl	816	6	1.2	377	3	US-09-041-889-29	Sequence 29, Appl
744	6	1.2	328	3	US-08-513-974B-371	Sequence 371, App	817	6	1.2	378	2	US-08-528-057-2	Sequence 2, Appl
745	6	1.2	329	4	US-08-513-974B-380	Sequence 380, App	818	6	1.2	378	2	US-08-244-205-7	Sequence 7, Appl
746	6	1.2	330	4	US-09-161-994A-6	Sequence 6, Appl	819	6	1.2	378	5	US-08-244-205-7	Sequence 7, Appl
747	6	1.2	331	2	US-08-907-674-1	Sequence 1, Appl	820	6	1.2	380	4	PCT-US92-10284-7	Sequence 7, Appl
748	6	1.2	331	2	US-08-907-674-1	Sequence 1, Appl	821	6	1.2	380	4	US-08-904-818-8	Sequence 8, Appl
749	6	1.2	331	3	US-09-391-959-1	Sequence 1, Appl	822	6	1.2	382	4	US-09-222-594-8	Sequence 8, Appl
750	6	1.2	332	5	US-09-391-959-1	Sequence 5, Appl	823	6	1.2	382	4	US-08-444-818-68	Sequence 68, Appl
751	6	1.2	332	5	PCT-US94-14277-2	Sequence 2, Appl	824	6	1.2	382	4	US-08-969-815-2	Sequence 2, Appl
752	6	1.2	333	4	US-09-556-877-292	Sequence 292, App	825	6	1.2	382	4	US-09-120-025-2	Sequence 2, Appl
753	6	1.2	333	4	US-09-620-412C-292	Sequence 292, App	826	6	1.2	384	2	US-09-710-481-2	Sequence 2, Appl
754	6	1.2	334	4	US-09-060-756-728	Sequence 728, App	827	6	1.2	384	2	US-08-401-068-2	Sequence 2, Appl
755	6	1.2	336	4	US-09-134-001C-5297	Sequence 5297, Ap	828	6	1.2	384	3	US-08-846-338-8	Sequence 8, Appl
756	6	1.2	337	2	US-09-013-634-2	Sequence 7, Appl	829	6	1.2	384	6	US-08-411-768B-7	Sequence 7, Appl
757	6	1.2	341	2	US-08-209-521-11	Sequence 11, Appl	830	6	1.2	389	2	US-08-139-195-2	Sequence 2, Appl
												US-08-605-106-13	Sequence 13, Appl

831	6	1.2	391	1	US-08-602-010A-6	Sequence 6, Appl1	904	6	1.2	458	1	US-08-336-618-24	Sequence 24, Appl1
832	6	1.2	391	1	US-08-680-726A-6	Sequence 6, Appl1	905	6	1.2	460	4	US-08-444-818-20	Sequence 20, Appl1
833	6	1.2	391	4	US-09-092-409-6	Sequence 6, Appl1	906	6	1.2	463	4	US-08-845-258-25	Sequence 25, Appl1
834	6	1.2	394	2	US-08-839-581A-4	Sequence 4, Appl1	907	6	1.2	463	4	US-08-990-571-25	Sequence 25, Appl1
835	6	1.2	394	4	US-09-023-591A-4	Sequence 4, Appl1	908	6	1.2	463	4	US-08-723-142A-25	Sequence 25, Appl1
836	6	1.2	397	1	US-07-956-697B-5	Sequence 5, Appl1	909	6	1.2	463	4	US-09-528-784A-25	Sequence 25, Appl1
837	6	1.2	397	1	US-08-263-098-5	Sequence 5, Appl1	910	6	1.2	464	2	US-08-602-725-12	Sequence 32, Appl1
838	6	1.2	397	1	US-08-434-255-2	Sequence 2, Appl1	911	6	1.2	464	4	US-08-970-134-3	Sequence 3, Appl1
839	6	1.2	397	1	US-08-434-255-4	Sequence 4, Appl1	912	6	1.2	465	4	US-08-030-369-1	Sequence 1, Appl1
840	6	1.2	397	1	US-08-459-867-2	Sequence 2, Appl1	913	6	1.2	465	4	US-09-482-971-1	Sequence 1, Appl1
841	6	1.2	397	1	US-08-459-867-4	Sequence 4, Appl1	914	6	1.2	465	4	US-09-134-001C-3856	Sequence 3856, Ap
842	6	1.2	397	1	US-08-460-327-2	Sequence 2, Appl1	915	6	1.2	466	4	US-08-970-134-1	Sequence 1, Appl1
843	6	1.2	397	1	US-08-460-327-4	Sequence 4, Appl1	916	6	1.2	466	4	US-08-914-375C-60	Sequence 60, Appl1
844	6	1.2	397	1	US-08-459-871-2	Sequence 2, Appl1	917	6	1.2	471	3	US-09-106-164-2	Sequence 2, Appl1
845	6	1.2	397	1	US-08-459-871-4	Sequence 4, Appl1	918	6	1.2	475	1	US-07-115-184-4	Sequence 4, Appl1
846	6	1.2	397	4	US-09-459-133-2	Sequence 2, Appl1	919	6	1.2	475	1	US-07-876-280-7	Sequence 7, Appl1
847	6	1.2	400	4	US-09-086-010-2	Sequence 2, Appl1	920	6	1.2	475	1	US-07-876-280-28	Sequence 28, Appl1
848	6	1.2	401	2	US-08-278-042C-2	Sequence 2, Appl1	921	6	1.2	475	1	US-07-935-310A-2	Sequence 2, Appl1
849	6	1.2	403	3	US-08-776-246-4	Sequence 4, Appl1	922	6	1.2	475	1	US-07-828-788A-2	Sequence 2, Appl1
850	6	1.2	404	2	US-08-244-205-9	Sequence 9, Appl1	923	6	1.2	475	1	US-08-049-783-4	Sequence 4, Appl1
851	6	1.2	404	5	PCT-US92-10284-9	Sequence 9, Appl1	924	6	1.2	475	1	US-08-147-189-2	Sequence 2, Appl1
852	6	1.2	405	4	US-09-144-914-5	Sequence 5, Appl1	925	6	1.2	475	1	US-08-316-301A-8	Sequence 8, Appl1
853	6	1.2	407	4	US-09-271-438A-4	Sequence 4, Appl1	926	6	1.2	475	2	US-08-904-278-2	Sequence 2, Appl1
854	6	1.2	407	4	US-09-271-438A-10	Sequence 10, Appl1	927	6	1.2	475	2	US-08-904-278-4	Sequence 4, Appl1
855	6	1.2	410	4	US-09-134-001C-4135	Sequence 4135, Ap	928	6	1.2	475	4	US-09-076-137-8	Sequence 8, Appl1
856	6	1.2	411	2	US-08-773-870-1	Sequence 1, Appl1	929	6	1.2	475	4	US-07-941-650A-6	Sequence 6, Appl1
857	6	1.2	412	4	US-09-423-439-10	Sequence 10, Appl1	930	6	1.2	475	4	US-09-222-594-2	Sequence 2, Appl1
858	6	1.2	413	1	US-08-295-676A-2	Sequence 2, Appl1	931	6	1.2	475	4	US-09-222-594-4	Sequence 4, Appl1
859	6	1.2	413	1	US-08-750-077-6	Sequence 6, Appl1	932	6	1.2	475	5	PCT-US92-03624-8	Sequence 8, Appl1
860	6	1.2	413	2	US-08-481-814A-8	Sequence 8, Appl1	933	6	1.2	475	5	PCT-US92-11337-2	Sequence 2, Appl1
861	6	1.2	413	3	US-08-836-582-2	Sequence 2, Appl1	934	6	1.2	475	5	PCT-US93-07409-2	Sequence 2, Appl1
862	6	1.2	413	3	US-08-948-591-2	Sequence 2, Appl1	935	6	1.2	485	1	US-07-881-075-1	Sequence 1, Appl1
863	6	1.2	413	4	US-09-265-566-2	Sequence 2, Appl1	936	6	1.2	485	1	US-08-120-827-1	Sequence 1, Appl1
864	6	1.2	413	4	US-09-242-737-4	Sequence 4, Appl1	937	6	1.2	485	1	US-08-478-612-1	Sequence 1, Appl1
865	6	1.2	416	4	US-09-330-611-8	Sequence 8, Appl1	938	6	1.2	485	4	US-09-384-212-2	Sequence 4, Appl1
866	6	1.2	421	4	US-09-332-478-6	Sequence 6, Appl1	939	6	1.2	491	4	US-09-262-856A-5	Sequence 5, Appl1
867	6	1.2	422	4	US-08-712-072C-2	Sequence 2, Appl1	940	6	1.2	491	4	US-09-134-001C-4612	Sequence 4612, Ap
868	6	1.2	424	4	US-09-120-817-2	Sequence 2, Appl1	941	6	1.2	492	4	US-08-134-001C-4847	Sequence 4847, Ap
869	6	1.2	427	4	US-09-199-737-4	Sequence 4, Appl1	942	6	1.2	493	1	US-08-362-512A-4	Sequence 4, Appl1
870	6	1.2	427	4	US-08-900-230-4	Sequence 4, Appl1	943	6	1.2	493	4	US-08-964-939-4	Sequence 4, Appl1
871	6	1.2	427	4	US-09-058-333A-4	Sequence 4, Appl1	944	6	1.2	494	3	US-08-993-260-3	Sequence 3, Appl1
872	6	1.2	428	2	US-08-717-587A-2	Sequence 2, Appl1	945	6	1.2	496	4	US-09-134-001C-3703	Sequence 3703, Ap
873	6	1.2	428	3	US-08-883-610A-2	Sequence 2, Appl1	946	6	1.2	497	4	US-09-594-193-2	Sequence 2, Appl1
874	6	1.2	428	3	US-09-118-319-5	Sequence 5, Appl1	947	6	1.2	498	4	US-09-232-468A-18	Sequence 18, Appl1
875	6	1.2	428	4	US-08-936-094A-2	Sequence 2, Appl1	948	6	1.2	498	4	US-09-232-468A-24	Sequence 24, Appl1
876	6	1.2	430	5	PCT-US93-04392-3	Sequence 3, Appl1	949	6	1.2	498	4	US-08-686-968C-231	Sequence 231, App
877	6	1.2	432	2	US-08-904-278-6	Sequence 6, Appl1	950	6	1.2	500	4	US-09-103-754A-5	Sequence 5, Appl1
878	6	1.2	432	4	US-09-222-594-6	Sequence 6, Appl1	951	6	1.2	505	3	US-09-326-203A-17	Sequence 17, Appl1
879	6	1.2	433	1	US-07-661-610C-2	Sequence 2, Appl1	952	6	1.2	505	3	US-08-303-861-2	Sequence 2, Appl1
880	6	1.2	433	4	US-09-364-230-14	Sequence 14, Appl1	953	6	1.2	507	4	US-09-091-097-10	Sequence 10, Appl1
881	6	1.2	433	4	US-09-134-001C-3085	Sequence 3085, Ap	954	6	1.2	507	4	US-08-091-097-34	Sequence 34, Appl1
882	6	1.2	434	2	US-08-710-249-4	Sequence 4, Appl1	955	6	1.2	508	2	US-08-724-281-1	Sequence 1, Appl1
883	6	1.2	434	4	US-09-220-157A-4	Sequence 4, Appl1	956	6	1.2	512	3	US-09-111-730-2	Sequence 2, Appl1
884	6	1.2	437	4	US-09-134-001C-3418	Sequence 3418, Ap	957	6	1.2	512	3	US-09-015-296-3	Sequence 3, Appl1
885	6	1.2	442	3	US-09-052-778-2	Sequence 2, Appl1	958	6	1.2	512	4	US-08-867-611-58	Sequence 58, Appl1
886	6	1.2	443	1	US-08-399-986B-2	Sequence 2, Appl1	959	6	1.2	512	4	US-08-724-224-4	Sequence 4, Appl1
887	6	1.2	443	2	US-08-493-754A-2	Sequence 2, Appl1	960	6	1.2	512	4	US-09-353-722-3	Sequence 3, Appl1
888	6	1.2	443	4	US-09-134-001C-4915	Sequence 415, Ap	961	6	1.2	513	1	US-08-464-266-2	Sequence 2, Appl1
889	6	1.2	444	5	US-09-252-292C-27	Sequence 27, Appl1	962	6	1.2	513	1	US-08-464-262-2	Sequence 2, Appl1
890	6	1.2	445	5	PCT-US94-05387-7	Sequence 7, Appl1	963	6	1.2	513	1	US-08-200-232-4	Sequence 4, Appl1
891	6	1.2	446	2	US-08-244-205-5	Sequence 5, Appl1	964	6	1.2	513	2	US-08-459-346-19	Sequence 19, Appl1
892	6	1.2	446	5	US-09-161-994A-5	Sequence 5, Appl1	965	6	1.2	513	2	US-07-989-847-8	Sequence 8, Appl1
893	6	1.2	446	5	PCT-US92-10284-5	Sequence 5, Appl1	966	6	1.2	513	3	US-08-889-419-19	Sequence 19, Appl1
894	6	1.2	446	5	PCT-US94-01321-10	Sequence 10, Appl1	967	6	1.2	513	4	US-08-459-411-8	Sequence 8, Appl1
895	6	1.2	450	2	US-08-861-464-2	Sequence 2, Appl1	968	6	1.2	513	4	US-08-464-514-2	Sequence 2, Appl1
896	6	1.2	450	2	US-08-396-001-2	Sequence 2, Appl1	969	6	1.2	513	4	US-08-486-403-2	Sequence 2, Appl1
897	6	1.2	450	4	US-09-323-433A-2	Sequence 2, Appl1	970	6	1.2	513	4	US-08-402-542-19	Sequence 19, Appl1
898	6	1.2	454	1	US-08-171-382-4	Sequence 4, Appl1	971	6	1.2	513	5	PCT-US93-07189-19	Sequence 19, Appl1
899	6	1.2	454	1	US-08-309-420-4	Sequence 4, Appl1	972	6	1.2	513	5	PCT-US93-02219-4	Sequence 4, Appl1
900	6	1.2	454	1	US-08-309-419-4	Sequence 4, Appl1	973	6	1.2	513	5	PCT-US95-02219A-4	Sequence 4, Appl1
901	6	1.2	454	4	US-09-294-531B-30	Sequence 30, Appl1	974	6	1.2	513	6	5187076-6	Patent No. 5187076
902	6	1.2	454	5	PCT-US95-11856-4	Sequence 4, Appl1	975	6	1.2	514	5	US-08-063-552-13	Sequence 13, Appl1
903	6	1.2	454	5	PCT-US95-11878-4	Sequence 4, Appl1	976	6	1.2	514	5	PCT-US93-05704-13	Sequence 13, Appl1

977	6	1.2	515	1	US-08-063-552-4	Sequence 4, Appli
978	6	1.2	515	5	PCT-US93-05704-4	Sequence 4, Appli
979	6	1.2	522	6	RE34606-6	Patent No. RE34,60
980	6	1.2	525	1	US-08-348-891A-2	Sequence 2, Appli
981	6	1.2	525	2	US-09-905-817-2	Sequence 2, Appli
982	6	1.2	525	4	US-09-351-224E-8	Sequence 8, Appli
983	6	1.2	526	2	US-08-853-659A-40	Sequence 40, Appli
984	6	1.2	529	1	US-08-548-509-2	Sequence 2, Appli
985	6	1.2	532	4	US-09-071-033-88	Sequence 88, Appli
986	6	1.2	535	1	US-07-872-644-49	Sequence 49, Appli
987	6	1.2	535	1	US-08-297-494-49	Sequence 49, Appli
988	6	1.2	535	1	US-08-297-510-49	Sequence 49, Appli
989	6	1.2	535	1	US-08-479-532-49	Sequence 49, Appli
990	6	1.2	535	1	US-08-455-526-49	Sequence 49, Appli
991	6	1.2	535	1	US-08-455-526-49	Sequence 49, Appli
992	6	1.2	535	3	US-08-813-574-2	Sequence 49, Appli
993	6	1.2	535	3	US-09-139-491-49	Sequence 49, Appli
994	6	1.2	535	5	PCT-US92-03222-49	Sequence 2, Appli
995	6	1.2	540	4	US-09-011-074-4	Sequence 49, Appli
996	6	1.2	542	4	US-09-380-420C-17	Sequence 17, Appli
997	6	1.2	545	4	US-09-351-224E-3	Sequence 4, Appli
998	6	1.2	550	1	US-08-484-493-2	Sequence 2, Appli
999	6	1.2	550	1	US-08-484-493-2	Sequence 2, Appli
1000	6	1.2	550	2	US-08-345-212-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1
US-08-809-326A-1
Sequence 1, Application US/08809326A

Patent No. 6165478
GENERAL INFORMATION:
APPLICANT: Izutsu, Hiroshi
APPLICANT: Obara, Kazuhiko
APPLICANT: Matsumoto, Akira
TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,
TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING
TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMANTS
TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF
TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT
TITLE OF INVENTION: OF ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT
TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND/OR
TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND
TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Penile & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,326A
FILING DATE: 19-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 224711/94
FILING DATE: 20-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106006/95
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106008/95
FILING DATE: 28-APR-1995

PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106009/95
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106010/95
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106011/95
FILING DATE: 28-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Miller, Charles E.
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 7426-043-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 488 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-809-326A-1

Query Match 98.2%; Score 487; DB 4; Length 488;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	4	MSISSSSGPDNKNIMSOVITSTPGVPOODKLSGNETOIOOTRGKTEMESDATTAG	63
DB	1	MSISSSSGPDNKNIMSOVITSTPGVPOODKLSGNETOIOOTRGKTEMESDATTAG	60
QY	64	ASGKDKTSTTTKETAPDQGVAAKESSESOKAGADTVGSAATTAATATKIMQTSI	123
DB	61	ASGKDKTSTTTKETAPDQGVAAKESSESOKAGADTVGSAATTAATATKIMQTSI	120
QY	124	EEASKSMESTLESLOSLSAOKKEVEAVVVAALSGKSSGSAKLEPPELPKPGVPRSEVI	183
DB	121	EEASKSMESTLESLOSLSAOKKEVEAVVVAALSGKSSGSAKLEPPELPKPGVPRSEVI	180
QY	184	EIGLALAKAIQTLGENTKSALSNVASTOAOQDTKLGLEKOKIIDEREYDOKRAE	243
DB	181	EIGLALAKAIQTLGENTKSALSNVASTOAOQDTKLGLEKOKIIDEREYDOKRAE	240
QY	244	OKSKDLEGMTDVTNVTMIAVSVAITVISIVAIFTCGAGLAGLAAGAAVGAAGA	303
DB	241	OKSKDLEGMTDVTNVTMIAVSVAITVISIVAIFTCGAGLAGLAAGAAVGAAGA	300
QY	304	AAATTVATQITVQAVVQAVKQAVITAVROAITAATKAIVKSGIKAFITVTKATAKAIK	363
DB	301	AAATTVATQITVQAVVQAVKQAVITAVROAITAATKAIVKSGIKAFITVTKATAKAIK	360
QY	364	GISKVPAGTQMTIAKPNFLSKVISLTSKWTYGVGVVVAAPALGKIMQMOQSEM	423
DB	361	GISKVPAGTQMTIAKPNFLSKVISLTSKWTYGVGVVVAAPALGKIMQMOQSEM	420
QY	424	VAQFOKEVGIQAAADMTSMFTQFMQASKTASQOTGSMNTOKAKRLGAIILKAYAI	483
DB	421	VAQFOKEVGIQAAADMTSMFTQFMQASKTASQOTGSMNTOKAKRLGAIILKAYAI	480
QY	484	SGAIGA 490	
DB	481	SGAIGA 487	

RESULT 2
US-08-809-326A-15
Sequence 15, Application US/08809326A
Patent No. 6165478
GENERAL INFORMATION:
APPLICANT: Izutsu, Hiroshi
APPLICANT: Obara, Kazuhiko

APPLICANT: Matsumoto, Akira
TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,
TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING
THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMANTS
TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF
ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT
TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA
PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND/OR
TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND
REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
PNEUMONIAE GENE
TITLE OF INVENTION:
TITLE OF INVENTION:
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,326A
FILING DATE: 19-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 224711/94
FILING DATE: 20-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106006/95
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106008/95
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106009/95
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106010/95
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106011/95
FILING DATE: 28-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Miller, Charles E.
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 7426-043-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 15:
SEQUENCE CHARACTERISTICS:
LENGTH: 649 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-809-326A-15

Query Match 98.28; Score 487; DB 4; Length 649;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MSISSSSGPNOKIMSOVLTSTPOGVPOODKLSGNETKOTIOROGKNTMESDATTAG 63
Db 162 MSISSSSGPNOKIMSOVLTSTPOGVPOODKLSGNETKOTIOROGKNTMESDATTAG 221

QY 64 ASGRDKTSSTTKETAPAOOGVAACKESSESOKAGADTGVSGAAATTTASNTATKIAMQTSI 123
Db 222 ASGRDKTSSTTKETAPAOOGVAACKESSESOKAGADTGVSGAAATTTASNTATKIAMQTSI 281

QY 124 EASKSMESTLESLOSISAQMKVEAVVAALSGKSSGSAKLEPPELPKRGVPRSPVI 183
Db 282 EASKSMESTLESLOSISAQMKVEAVVAALSGKSSGSAKLEPPELPKRGVPRSPVI 341

QY 184 EIGALAKAIIOTLEATFSALSNASTOQADQNTKLGLEQAIKIDEREYOEKAAE 243
Db 342 EIGALAKAIIOTLEATFSALSNASTOQADQNTKLGLEQAIKIDEREYOEKAAE 401

QY 244 OKSKDLEGTMTVTMTVIAVSVAITVISVAIIFTCGAGLAGLAAGAAVAAAAGAGA 303
Db 402 OKSKDLEGTMTVTMTVIAVSVAITVISVAIIFTCGAGLAGLAAGAAVAAAAGAGA 461

QY 304 AAATTTATQTTVOAVQAVKAVITAVQAQITTAIKAIVKSGIAFTITVKATAIKAIK 363
Db 462 AAATTTATQTTVOAVQAVKAVITAVQAQITTAIKAIVKSGIAFTITVKATAIKAIK 521

QY 364 GISKVFAGTOMIAKNPFLSKVISLSLTSKWVYGVVVAAPALGKIMOMOLEMOON 423
Db 522 GISKVFAGTOMIAKNPFLSKVISLSLTSKWVYGVVVAAPALGKIMOMOLEMOON 581

QY 424 VAOFQEVGKLOAAADMI SMETOFMOQASKIASKOTGESNEMTQATKGAQILKAYAI 483
Db 582 VAOFQEVGKLOAAADMI SMETOFMOQASKIASKOTGESNEMTQATKGAQILKAYAI 641

QY 484 SGATAGA 490
Db 642 SGATAGA 648

RESULT 3
US-08-809-326A-5
Sequence 5, Application US/08809326A
Patent No. 6165478
GENERAL INFORMATION:
APPLICANT: Izutsu, Hiroshi
APPLICANT: Obara, Kazuhiko
APPLICANT: Matsumoto, Akira
TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES, DNAS CODING
THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMANTS
CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF
ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT
OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA
PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND
TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND
REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
PNEUMONIAE GENE
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,326A
FILING DATE: 19-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 224711/94
FILING DATE: 20-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106006/95
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106008/95

```

? FILING DATE: 28-APR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 106009/95
? FILING DATE: 28-APR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 106010/95
? FILING DATE: 28-APR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 106011/95
? FILING DATE: 28-APR-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: Miller, Charles E.
? REGISTRATION NUMBER: 24,576
? REFERENCE/DOCKET NUMBER: 7426-043-999
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 790-9090
? TELEFAX: (212) 869-8864/9741
? TELEX: 66141 PENNIE
? INFORMATION FOR SEQ ID NO: 5:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 259 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
US-08-809-326A-5

```

```

Query Match          52.2%; Score 259; DB 4; Length 259;
Best Local Similarity 100.0%; Pred. No. 1.5e-226;
Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 MSISSSSGPNOKNIMSOVLSTPQGVPOODKLSGNETKQIOOTRGKNTMESDATING 63
DB 1 MSISSSSGPNOKNIMSOVLSTPQGVPOODKLSGNETKQIOOTRGKNTMESDATING 60
QY 64 ASGKDKTSSTTKETAPQGVAAAGKSSSESOKAGADTVGSAATTAATSNATKIMQTSI 123
DB 61 ASGKDKTSSTTKETAPQGVAAAGKSSSESOKAGADTVGSAATTAATSNATKIMQTSI 120
QY 124 EASKSMESTLESLSLSAOMKEVEAVVAALSGKSSGSASAKLETPELPKPGVTPRSEVI 183
DB 121 EASKSMESTLESLSLSAOMKEVEAVVAALSGKSSGSASAKLETPELPKPGVTPRSEVI 180
QY 184 EIGLALAKAIQTLGEATKSLSNYASTQAOADOTNKLGLEKQAIKIDKEREYQEMKAAE 243
DB 181 EIGLALAKAIQTLGEATKSLSNYASTQAOADOTNKLGLEKQAIKIDKEREYQEMKAAE 240
QY 244 QKSKDEGTMDTVNTVYMA 262
DB 241 QKSKDEGTMDTVNTVYMA 259

```

```

RESULT 4
US-08-809-326A-2
? Sequence 2, Application US/08809326A
? Patent No. 6165478
? GENERAL INFORMATION:
? APPLICANT: Izutsu, Hiroshi
? APPLICANT: Obara, Kazuhiko
? APPLICANT: Matsumoto, Akira
? TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,
? TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING
? TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMANTS
? TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF
? TITLE OF INVENTION: OF ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT
? TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA
? TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND/OR
? TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
? NUMBER OF SEQUENCES: 31
? CORRESPONDENCE ADDRESS:
? ADDRESSEE: Pennie & Edmonds LLP
? STREET: 1155 Avenue of the Americas

```

```

? CITY: New York
? STATE: New York
? COUNTRY: USA
? ZIP: 10036-2711
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent In Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/08/809,326A
? FILING DATE: 19-MAR-1997
? CLASSIFICATION: 435
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 224711/94
? FILING DATE: 20-SEP-1994
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 106006/95
? FILING DATE: 28-APR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 106008/95
? FILING DATE: 28-APR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 106009/95
? FILING DATE: 28-APR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 106010/95
? FILING DATE: 28-APR-1995
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 106011/95
? FILING DATE: 28-APR-1995
? ATTORNEY/AGENT INFORMATION:
? NAME: Miller, Charles E.
? REGISTRATION NUMBER: 24,576
? REFERENCE/DOCKET NUMBER: 7426-043-999
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (212) 790-9090
? TELEFAX: (212) 869-8864/9741
? TELEX: 66141 PENNIE
? INFORMATION FOR SEQ ID NO: 2:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 271 amino acids
? TYPE: amino acid
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
US-08-809-326A-2

```

```

Query Match          52.2%; Score 259; DB 4; Length 271;
Best Local Similarity 100.0%; Pred. No. 1.6e-226;
Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 4 MSISSSSGPNOKNIMSOVLSTPQGVPOODKLSGNETKQIOOTRGKNTMESDATING 63
DB 1 MSISSSSGPNOKNIMSOVLSTPQGVPOODKLSGNETKQIOOTRGKNTMESDATING 60
QY 64 ASGKDKTSSTTKETAPQGVAAAGKSSSESOKAGADTVGSAATTAATSNATKIMQTSI 123
DB 61 ASGKDKTSSTTKETAPQGVAAAGKSSSESOKAGADTVGSAATTAATSNATKIMQTSI 120
QY 124 EASKSMESTLESLSLSAOMKEVEAVVAALSGKSSGSASAKLETPELPKPGVTPRSEVI 183
DB 121 EASKSMESTLESLSLSAOMKEVEAVVAALSGKSSGSASAKLETPELPKPGVTPRSEVI 180
QY 184 EIGLALAKAIQTLGEATKSLSNYASTQAOADOTNKLGLEKQAIKIDKEREYQEMKAAE 243
DB 181 EIGLALAKAIQTLGEATKSLSNYASTQAOADOTNKLGLEKQAIKIDKEREYQEMKAAE 240
QY 244 QKSKDEGTMDTVNTVYMA 262
DB 241 QKSKDEGTMDTVNTVYMA 259

```

RESULT 5

US-08-809-326A-16
 ; Sequence 16, Application US/08809326A
 ; Patent No. 6165478
 ; GENERAL INFORMATION:
 ; APPLICANT: Izutsu, Hiroshi
 ; APPLICANT: Obara, Kazuhiko
 ; APPLICANT: Matsumoto, Akira
 ; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,
 ; TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING
 ; THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMANTS
 ; TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF
 ; TITLE OF INVENTION: ANTI-BODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT
 ; TITLE OF INVENTION: OF ANTI-BODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA
 ; TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND/OR
 ; TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND
 ; REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
 ; TITLE OF INVENTION: PNEUMONIAE GENE
 ; NUMBER OF SEQUENCES: 31
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds LLP
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/809,326A
 ; FILING DATE: 19-MAR-1997
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 224711/94
 ; FILING DATE: 20-SEP-1994
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 106006/95
 ; FILING DATE: 28-APR-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 106008/95
 ; FILING DATE: 28-APR-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 106009/95
 ; FILING DATE: 28-APR-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 106010/95
 ; FILING DATE: 28-APR-1995
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: JP 106011/95
 ; FILING DATE: 28-APR-1995
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Miller, Charles E.
 ; REGISTRATION NUMBER: 24,576
 ; REFERENCE/DOCKET NUMBER: 7426-043-999
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-8864/9741
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 16:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 432 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: peptide
 ; US-08-809-326A-16

Query Match 52.2%; Score 259; DB 4; Length 432;
 Best Local Similarity 100.0%; Pred. No. 2.5e-226;
 Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

4 MSISSSGPDNKNIMQVLTSTPGVPGQDKLSGNETKQIQTRGKNTMESDATTAG 63

Db 162 MSISSSGPDNKNIMQVLTSTPGVPGQDKLSGNETKQIQTRGKNTMESDATTAG 221
 Qy 64 ASGDKDTSSSTTKTETAPQOQVAAKESSESQKAGADTGVGAATPASTATKIAMQTSI 123
 Db 222 ASGDKDTSSSTTKTETAPQOQVAAKESSESQKAGADTGVGAATPASTATKIAMQTSI 281
 Qy 124 EEAQSMSTLESQSLAQMKEVAVVAALSGSSGSAKLEPPELPKGVTPRSEVI 183
 Db 282 EEAQSMSTLESQSLAQMKEVAVVAALSGSSGSAKLEPPELPKGVTPRSEVI 341
 Qy 184 EIGLALAKIOTLGEATRSALSNASTOAOADOTNKLGLERQAIKIDEREYQEMKAAE 243
 Db 342 EIGLALAKIOTLGEATRSALSNASTOAOADOTNKLGLERQAIKIDEREYQEMKAAE 401
 Qy 244 QKSKDLEGTMDTVMTVMA 262
 Db 402 QKSKDLEGTMDTVMTVMA 420

RESULT 6
 ; US-09-135-994-11
 ; Sequence 11, Application US/09135994A
 ; Patent No. 6280938
 ; GENERAL INFORMATION:
 ; APPLICANT: Rannum et al.
 ; TITLE OF INVENTION: SCAT GENE AND METHODS OF USE
 ; FILE REFERENCE: University of Minnesota
 ; CURRENT APPLICATION NUMBER: US/09/135,994A
 ; CURRENT FILING DATE: 1998-08-18
 ; EARLIER APPLICATION NUMBER: 60/056,170
 ; EARLIER FILING DATE: 1997-08-19
 ; NUMBER OF SEQ ID NOS: 14
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 11
 ; LENGTH: 27
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-09-135-994-11

Query Match 1.6%; Score 8; DB 4; Length 27;
 Best Local Similarity 100.0%; Pred. No. 2.6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 294 AAAGGAA 301
 Db 17 AAAGGAA 24

RESULT 7
 ; US-08-425-069-50
 ; Sequence 50, Application US/08425069
 ; Patent No. 5728810
 ; GENERAL INFORMATION:
 ; APPLICANT: Lewis, Randolph V.
 ; APPLICANT: Xu, Ming
 ; APPLICANT: Himman, Michael B.
 ; TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
 ; TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
 ; TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
 ; NUMBER OF SEQUENCES: 69
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Birch, Stewart, Kolasch & Birch
 ; STREET: 301 No. 5728810th Washington Street
 ; CITY: Falls Church
 ; STATE: Virginia
 ; COUNTRY: U.S.A.
 ; ZIP: 22046
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25

```

CURRENT APPLICATION DATA:
  APPLICATION NUMBER: US/08/425.069
  FILING DATE: 19-APR-1995
  CLASSIFICATION: 435
  ATTORNEY/AGENT INFORMATION:
    NAME: Murphy Jr., Gerald M
    REGISTRATION NUMBER: 28,977
    REFERENCE/DOCKET NUMBER: 1447-106P
  TELECOMMUNICATION INFORMATION:
    TELEPHONE: (703) 205-8000
    TELEFAX: (703) 205-8050
  TELEX:
  INFORMATION FOR SEQ ID NO: 50:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 28 amino acids
      TYPE: amino acid
      STRANDEDNESS: not relevant
      TOPOLOGY: not relevant
    MOLECULE TYPE: peptide
    HYPOTHEetical: NO
    FRAGMENT TYPE: Internal
    ORIGINAL SOURCE:
      ORGANISM: nephila clavipes
  FEATURE:
    NAME/KEY: Peptide
    LOCATION: 1..28
    OTHER INFORMATION: /label= silk1_repeat
US-08-425-069-50

Query Match      1.6%; Score 8; DB 1; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGA 300
   |||||
Db 12 GAAAGGA 19

RESULT 8
US-08-317-844B-50
  Sequence 50, Application US/08317844B
  Patent No. 5989894
  GENERAL INFORMATION:
    APPLICANT: Lewis, Randolph V.
    APPLICANT: Xu, Ming
    TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
    TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
    TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
    NUMBER OF SEQUENCES: 62
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Birch, Stewart, Kolasch & Birch
    STREET: 301 No. 5989894th Washington Street
    CITY: Falls Church
    STATE: Virginia
    COUNTRY: U.S.A.
    ZIP: 22046
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patent Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/317,844B
    FILING DATE: 04-OCT-1994
    CLASSIFICATION: 435
    ATTORNEY/AGENT INFORMATION:
      NAME: Murphy Jr., Gerald M
      REGISTRATION NUMBER: 28,977
      REFERENCE/DOCKET NUMBER: 1447-105P
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: (703) 241-1300
      TELEFAX: (703) 241-2848
```

```

TELEX: 248345
  INFORMATION FOR SEQ ID NO: 50:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 28 amino acids
      TYPE: amino acid
      STRANDEDNESS: not relevant
      TOPOLOGY: not relevant
    MOLECULE TYPE: peptide
    HYPOTHEtical: NO
    FRAGMENT TYPE: Internal
    ORIGINAL SOURCE:
      ORGANISM: nephila clavipes
  FEATURE:
    NAME/KEY: Peptide
    LOCATION: 1..28
    OTHER INFORMATION: /label= silk1_repeat
US-08-317-844B-50

Query Match      1.6%; Score 8; DB 2; Length 28;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGA 300
   |||||
Db 12 GAAAGGA 19

RESULT 9
US-08-425-069-61
  Sequence 61, Application US/08425069
  Patent No. 5728810
  GENERAL INFORMATION:
    APPLICANT: Lewis, Randolph V.
    APPLICANT: Xu, Ming
    TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
    TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
    TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
    NUMBER OF SEQUENCES: 69
  CORRESPONDENCE ADDRESS:
    ADDRESSEE: Birch, Stewart, Kolasch & Birch
    STREET: 301 No. 5728810th Washington Street
    CITY: Falls Church
    STATE: Virginia
    COUNTRY: U.S.A.
    ZIP: 22046
  COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patent Release #1.0, Version #1.25
  CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/425.069
    FILING DATE: 19-APR-1995
    CLASSIFICATION: 435
    ATTORNEY/AGENT INFORMATION:
      NAME: Murphy Jr., Gerald M
      REGISTRATION NUMBER: 28,977
      REFERENCE/DOCKET NUMBER: 1447-106P
    TELECOMMUNICATION INFORMATION:
      TELEPHONE: (703) 205-8000
      TELEFAX: (703) 205-8050
    TELEX:
  INFORMATION FOR SEQ ID NO: 61:
    SEQUENCE CHARACTERISTICS:
      LENGTH: 30 amino acids
      TYPE: amino acid
      STRANDEDNESS: not relevant
      TOPOLOGY: not relevant
    MOLECULE TYPE: peptide
    HYPOTHEtical: NO
    FRAGMENT TYPE: Internal
    ORIGINAL SOURCE:
```

ORGANISM: nephila clavipes
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..30
OTHER INFORMATION: /label= silk1_repeat
US-08-425-069-61

Query Match 1.6%; Score 8; DB 1;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGCA 300
|||||||
DB 12 GAAAGCA 19

RESULT 10
US-08-317-844B-61
Sequence 61, Application US/08317844B
Patent No. 5989894
GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V.
APPLICANT: Xu, Ming
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 301 No. 5989894th Washington Street
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22046

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/317,844B
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-105P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 241-1300
TELEFAX: (703) 241-2848
TELEX: 248345

INFORMATION FOR SEQ ID NO: 61:
SEQUENCE CHARACTERISTICS:
LENGTH: 30 amino acids
TYPE: amino acid
STRANDEDNESS: not relevant
TOPOLOGY: not relevant
MOLECULE TYPE: peptide
HYPOTHETICAL: NO
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
ORGANISM: nephila clavipes
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..30
OTHER INFORMATION: /label= silk1_repeat
US-08-317-844B-61

Query Match 1.6%; Score 8; DB 2;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGCA 300
|||||||
DB 12 GAAAGCA 19

RESULT 11
US-08-425-069-6
Sequence 6, Application US/08425069
Patent No. 5728810
GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V.
APPLICANT: Xu, Ming
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
NUMBER OF SEQUENCES: 69
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 301 No. 5728810th Washington Street
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22046

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/425,069
FILING DATE: 19-APR-1995
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-106P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX:

INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 33 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: Internal
ORIGINAL SOURCE:
ORGANISM: Nephila clavipes
FEATURE:
NAME/KEY: Peptide
LOCATION: 1..33
OTHER INFORMATION: /label= representative
OTHER INFORMATION: /note="this peptide is a representative one that
OTHER INFORMATION: illustrates the g9x9yg hexamer repeat motif of the
OTHER INFORMATION: spider silk protein I."

US-08-425-069-6

Query Match 1.6%; Score 8; DB 1;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGCA 300
|||||||
DB 6 GAAAGCA 13

RESULT 12
US-08-317-844B-6
Sequence 6, Application US/08317844B
Patent No. 5989894
GENERAL INFORMATION:

```

; APPLICANT: Lewis, Randolph V.
; APPLICANT: Xu, Ming
; APPLICANT: Hinman, Michael B.
; TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
; TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
; TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 No. 5989894th Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/317,844B
; FILING DATE: 04-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1447-105P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 241-1300
; TELEFAX: (703) 241-2848
; TELEX: 248345
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 33 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: internal
; ORIGINAL SOURCE:
; ORGANISM: Nephila clavipes
; FEATURE:
; NAME/KEY: Peptide
; LOCATION: 1..33
; OTHER INFORMATION: /label=representative
; OTHER INFORMATION: /note="this peptide is a representative one that
; OTHER INFORMATION: illustrates the 99x99g hexamer repeat motif of the
; OTHER INFORMATION: spider silk protein I."
; US-08-317-844B-6
;
; Query Match 1.6%; Score 8; DB 2; Length 33;
; Best Local Similarity 100.0%; Pred. No. 3.2;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 293 GAAAGGA 300
; DB 6 GAAAGGA 13
;
; RESULT 13
; US-08-185-414E-2
; Sequence 2, Application US/08185414E
; Patent No. 5556953
; GENERAL INFORMATION:
; APPLICANT: Zhang, Lei
; APPLICANT: Vijay, Hari M.
; APPLICANT: Rode, Harold
; TITLE OF INVENTION: ALLERGEN OF CLADOSPORIUM HERBARUM
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: George A. Seaby
; ADDRESS: Seaby and Maclean
; STREET: 880 Wellington Street, Suite 708
; CITY: Ottawa

```

```

; COUNTRY: Canada
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,414E
; FILING DATE: January 24, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: George A. Seaby
; REGISTRATION NUMBER: 24,034
; REFERENCE/DOCKET NUMBER: 1747
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 232-5815
; TELEFAX: (613) 232-5831
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 111 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; FEATURE:
; NAME/KEY: Ch2.1 protein
; LOCATION: 1..111
; US-08-185-414E-2

```

```

; Query Match 1.6%; Score 8; DB 1; Length 111;
; Best Local Similarity 100.0%; Pred. No. 11;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 293 GAAAGGA 300
; DB 78 GAAAGGA 85

```

```

; RESULT 14
; US-09-135-994-12
; Sequence 12, Application US/09135994A
; Patent No. 6280938
; GENERAL INFORMATION:
; APPLICANT: Rannum et al.
; TITLE OF INVENTION: SCAT GENE AND METHODS OF USE
; FILE REFERENCE: University of Minnesota
; CURRENT APPLICATION NUMBER: US/09/135,994A
; CURRENT FILING DATE: 1998-08-18
; EARLIER APPLICATION NUMBER: 60/056,170
; EARLIER FILING DATE: 1997-08-19
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO: 12
; LENGTH: 129
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-135-994-12

```

```

; Query Match 1.6%; Score 8; DB 4; Length 129;
; Best Local Similarity 100.0%; Pred. No. 12;
; Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
;
; QY 294 AAAAGGA 301
; DB 17 AAAAGGA 24

```

```

; RESULT 15
; US-08-911-319A-1
; Sequence 1, Application US/08911319A
; Patent No. 5968798
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.

```

APPLICANT: Corley, Neil C.
APPLICANT: Shah, Puri
TITLE OF INVENTION: HUMAN GLUTAREDOXIN BETA
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTED for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/911,319A
FILING DATE: August 14, 1997
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Muenzen, Colette C.
REGISTRATION NUMBER: 39,784
REFERENCE/DOCKET NUMBER: PF-0363 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 164 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: THPINOT03
CLONE: 2447829
US-08-911-319A-1

Query Match 1.6%; Score 8; DB 2; Length 164;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 299 GAAGAAA 306
Db 29 GAAGAAA 36

RESULT 16
US-09-352-619-1
Sequence 1, Application US/09352619
Patent No. 6084070
GENERAL INFORMATION:
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
TITLE OF INVENTION: HUMAN GLUTAREDOXIN BETA
NUMBER OF SEQUENCES: 6
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Dr.
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTED for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/352,619

FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/911,319
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Muenzen, Colette C.
REGISTRATION NUMBER: 39,784
REFERENCE/DOCKET NUMBER: PF-0363 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-855-0555
TELEFAX: 650-845-4166
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 164 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: THPINOT03
CLONE: 2447829
US-09-352-619-1

Query Match 1.6%; Score 8; DB 3; Length 164;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 299 GAAGAAA 306
Db 29 GAAGAAA 36

RESULT 17
US-08-775-009-37
Sequence 37, Application US/08775009
Patent No. 5935783
GENERAL INFORMATION:
APPLICANT: Gong, Weijong
APPLICANT: Emanuel, Beverly S.
APPLICANT: Budarf, Marcia L.
TITLE OF INVENTION: No. 5935783el Genes Mapping in the Digestion and
TITLE OF INVENTION: Velocardiolfacial Syndrome Minimal Critical Region
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz &
ADDRESSEE: No. 5935783rls, LLP
STREET: One Liberty Place - 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: U.S.A.
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/775,009
FILING DATE: 27-DEC-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Trujillo, Doreen Yanko
REGISTRATION NUMBER: 35,719
REFERENCE/DOCKET NUMBER: CH-0681
TELECOMMUNICATION INFORMATION:
TELEPHONE: (215) 568-3100
TELEFAX: (215) 568-3439
INFORMATION FOR SEQ ID NO: 37:
SEQUENCE CHARACTERISTICS:
LENGTH: 205 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein

US-08-775-009-37

Query Match

1.6%; Score 8; DB 2; Length 205;
Best Local Similarity 100.0%; Pred. No. 19;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 AAAAGGA 301

Db 2 AAAAGGA 9

RESULT 18

US-08-364-081-3

Sequence 3; Application US/08364081

Patent No. 5585464

GENERAL INFORMATION:

APPLICANT: Ramesh K. Prakash

TITLE OF INVENTION: Recombinant Antigen for Diagnosing

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESS: Thorpe, No. 5585464th & Western

STREET: 9035 South 700 East, Suite 200

CITY: Sandy

STATE: Utah

COUNTRY: USA

ZIP: 84070

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb storage

COMPUTER: AST Advantage NB-SX20

OPERATING SYSTEM: DOS 6.2

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/364,081

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/019,780

FILING DATE: 19-FEB-1993

ATTORNEY/AGENT INFORMATION:

NAME: Alan J. Howarth

REGISTRATION NUMBER: 36,553

REFERENCE/DOCKET NUMBER: T781CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (801)566-6633

TELEFAX: (801)566-0750

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 331 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-364-081-3

Query Match

Best Local Similarity 100.0%; Score 8; DB 1; Length 331;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGA 300

Db 260 GAAAGGA 267

RESULT 19

US-08-630-552-3

Sequence 3; Application US/08630552

Patent No. 5723314

GENERAL INFORMATION:

APPLICANT: Ramesh K. Prakash

TITLE OF INVENTION: Recombinant Antigen for Diagnosing

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESS: Thorpe, No. 5723314th & Western

STREET: 9035 South 700 East, Suite 200

CITY: Sandy

STATE: Utah

COUNTRY: USA

ZIP: 84070

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb storage

COMPUTER: AST Advantage NB-SX20

OPERATING SYSTEM: DOS 6.2

SOFTWARE: Word Perfect 5.1

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/630,552

FILING DATE: 10-APR-1996

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/364,081

FILING DATE:

APPLICATION NUMBER: 08/019,780

FILING DATE: 19-FEB-1993

ATTORNEY/AGENT INFORMATION:

NAME: Alan J. Howarth

REGISTRATION NUMBER: 36,553

REFERENCE/DOCKET NUMBER: T781CIP

TELECOMMUNICATION INFORMATION:

TELEPHONE: (801)566-6633

TELEFAX: (801)566-0750

INFORMATION FOR SEQ ID NO: 3:

SEQUENCE CHARACTERISTICS:

LENGTH: 331 amino acids

TYPE: amino acid

TOPOLOGY: linear

US-08-630-552-3

Query Match

Best Local Similarity 100.0%; Score 8; DB 1; Length 331;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGA 300

Db 260 GAAAGGA 267

RESULT 20

PCT-US95-16558-3

Sequence 3; Application PC/TUS9516558

GENERAL INFORMATION:

APPLICANT: Ramesh K. Prakash

TITLE OF INVENTION: Recombinant Antigen for Diagnosing

NUMBER OF SEQUENCES: 3

CORRESPONDENCE ADDRESS:

ADDRESS: Thorpe, North & Western

STREET: 9035 South 700 East, Suite 200

CITY: Sandy

STATE: Utah

COUNTRY: USA

ZIP: 84070

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette, 3.5 inch, 720 Kb

COMPUTER: IBM ThinkPad 340

OPERATING SYSTEM: DOS 6.2

SOFTWARE: Word Perfect 6.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: PCT/US95/16558

FILING DATE: 27-DEC-1994

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/019,780

FILING DATE: 19-FEB-1993

ATTORNEY/AGENT INFORMATION:

NAME: Alan J. Howarth

REGISTRATION NUMBER: 36,553

REFERENCE/DOCKET NUMBER: T781CIP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (801)566-6633
TELEFAX: (801)566-0750
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 331 amino acids
TYPE: amino acid
TOPOLOGY: linear
PCT-US95-16558-3

Query Match 1.6%; Score 8; DB 5; Length 331;
Best Local Similarity 100.0%; Pred. No. 31;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 293 GAAAGGA 300
Db 260 GAAAGGA 267

RESULT 21
US-09-347-801-18
Sequence 18, Application US/09347801
Patent No. 6262345
GENERAL INFORMATION:
APPLICANT: Allen, Steve
APPLICANT: Lee, Jian Ming
TITLE OF INVENTION: Plant Protein Kinases
FILE REFERENCE: BB-1171
CURRENT APPLICATION NUMBER: US/09/347,801
CURRENT FILING DATE: 1999-07-02
EARLIER APPLICATION NUMBER: 60/092,438
EARLIER FILING DATE: July 10, 1998
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Microsoft Office 97
SEQ ID NO 18
LENGTH: 625
TYPE: PRT
ORGANISM: Zea mays
US-09-347-801-18

Query Match 1.6%; Score 8; DB 4; Length 625;
Best Local Similarity 100.0%; Pred. No. 58;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 288 AGAAGGA 295
Db 144 AGAAGGA 151

RESULT 22
US-08-556-978B-19
Sequence 19, Application US/08556978B
Patent No. 6268169
GENERAL INFORMATION:
APPLICANT: FAHNESTOCK, STEPHEN F.
TITLE OF INVENTION: NOVEL RECOMBINANTLY PRODUCED
TITLE OF INVENTION: SPIDER SILK ANALOGS
NUMBER OF SEQUENCES: 107
CORRESPONDENCE ADDRESS:
ADDRESSEE: E. I. DU PONT DE NEMOURS AND COMPANY
STREET: 1007 MARKET STREET
CITY: WILMINGTON
STATE: DELAWARE
COUNTRY: UNITED STATES OF AMERICA
ZIP: 19898
COMPUTER READABLE FORM:
MEDIUM TYPE: DISKETTE, 3.50 INCH
COMPUTER: IBM PC COMPATIBLE
OPERATING SYSTEM: MICROSOFT WINDOWS 95
SOFTWARE: MICROSOFT WORD FOR WINDOWS 95
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/556,978B

FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/077,600
FILING DATE: JUNE 15, 1993
ATTORNEY/AGENT INFORMATION:
NAME: FLOYD, LINDA AXAMETHY
REGISTRATION NUMBER: 33,692
REFERENCE/DOCKET NUMBER: CR-9389-A
TELECOMMUNICATION INFORMATION:
TELEPHONE: 302-892-8112
TELEFAX: 302-773-0164
INFORMATION FOR SEQ ID NO: 19:
SEQUENCE CHARACTERISTICS:
LENGTH: 651 amino acids
TYPE: amino acid
STRANDEDNESS: unknown
TOPOLOGY: unknown
MOLECULE TYPE: protein
US-08-556-978B-19

Query Match 1.6%; Score 8; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 293 GAAAGGA 300
Db 317 GAAAGGA 324

RESULT 23
US-09-247-806-1
Sequence 1, Application US/09247806
Patent No. 6280747
GENERAL INFORMATION:
APPLICANT: PHILIPPE, Michel
APPLICANT: GARSON, Jean-Claude
TITLE OF INVENTION: ARRANDEAU, Jean-Pierre
TITLE OF INVENTION: COSMETIC OR DERMATOLOGICAL COMPOSITION CONTACTING AT
TITLE OF INVENTION: LEAST ONE NATURAL OR RECOMBINANT SPIDER SILK OR AN
FILE REFERENCE: 6388-0365-0
CURRENT APPLICATION NUMBER: US/09/247,806
CURRENT FILING DATE: 1999-02-11
EARLIER APPLICATION NUMBER: FR 98/01614
EARLIER FILING DATE: 1998-02-11
NUMBER OF SEQ ID NOS: 14
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1
LENGTH: 651
TYPE: PRT
ORGANISM: Nephila clavipes
US-09-247-806-1

Query Match 1.6%; Score 8; DB 4; Length 651;
Best Local Similarity 100.0%; Pred. No. 60;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 293 GAAAGGA 300
Db 317 GAAAGGA 324

RESULT 24
US-08-425-069-2
Sequence 2, Application US/08425069
Patent No. 5728610
GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V.
APPLICANT: Xu, Ming
APPLICANT: Himan, Michael B.
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL

;; TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
;; NUMBER OF SEQUENCES: 69
;; CORRESPONDENCE ADDRESS:
;; ADDRESS: Birch, Stewart, Kolasch & Birch
;; STREET: 301 No. 5728810th Washington Street
;; CITY: Falls Church
;; STATE: Virginia
;; COUNTRY: U.S.A.
;; ZIP: 22046
;;
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: Patent Release #1.0, Version #1.25
;;
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/425,069
;; FILING DATE: 19-APR-1995
;; CLASSIFICATION: 435
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Murphy Jr., Gerald M
;; REGISTRATION NUMBER: 28,977
;; REFERENCE/DOCKET NUMBER: 1447-106P
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 205-8000
;; TELEFAX: (703) 205-8050
;; TELEX:
;;
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 718 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
US-08-425-069-2

Query Match
Best Local Similarity 1.6%; Score 8; DB 1; Length 718;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGA 300
|||
Db 317 GAAAGGA 324

RESULT 25
US-08-317-844B-2
; Sequence 2, Application US/08317844B
; Patent No. 5989894
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V.
; APPLICANT: Xu, Ming
; APPLICANT: Himan, Michael B.
; TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
; TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
; TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
; NUMBER OF SEQUENCES: 62
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 No. 5989894th Washington Street
; CITY: Falls Church
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/317,844B
; FILING DATE: 04-OCT-1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M

;; REGISTRATION NUMBER: 28,977
;; REFERENCE/DOCKET NUMBER: 1447-105P
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: (703) 241-1300
;; TELEFAX: (703) 241-2848
;; TELEX: 248345
;;
;; INFORMATION FOR SEQ ID NO: 2:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 718 amino acids
;; TYPE: amino acid
;; TOPOLOGY: linear
;; MOLECULE TYPE: protein
;;
US-08-317-844B-2

Query Match
Best Local Similarity 1.6%; Score 8; DB 2; Length 718;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGA 300
|||
Db 317 GAAAGGA 324

RESULT 26
US-09-034-177-3
; Sequence 3, Application US/09034177
; Patent No. 6127146
; GENERAL INFORMATION:
; APPLICANT: Lal, Preeti
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; TITLE OF INVENTION: HUMAN FIBROUS PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/034,177
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION NUMBER:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0486 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 747 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: Genbank
; CLONE: GI 1174414
; US-09-034-177-3

Query Match
Best Local Similarity 1.6%; Score 8; DB 3; Length 747;
Matches 100.0%; Pred. No. 69;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGCA 300
|||||
Db 624 GAAAGCA 631

RESULT 27
US-08-630-915A-37

; Sequence 37, Application US/08630915A
; Patent No. 6309820

; GENERAL INFORMATION:

; APPLICANT: SPARKS, Andrew B.

; APPLICANT: HOFFMAN, No. 6309820h

; APPLICANT: KAY, Brian K.

; APPLICANT: FOWLES, Dana M.

; TITLE OF INVENTION: POLYPEPTIDES HAVING A FUNCTIONAL

; TITLE OF INVENTION: DOMAIN OF INTEREST AND METHODS OF IDENTIFYING AND

; NUMBER OF SEQUENCES: 227

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Pennie & Edmonds LLP

; STREET: 1155 Avenue of the Americas

; City: New York

; STATE: New York

; COUNTRY: USA

; ZIP: 10036-2711

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: IBM PC compatible

; SOFTWARE: Patentin Release #1.0, Version #1.30

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/630,915A

; FILING DATE: 03-APR-1996

; CLASSIFICATION: 536

; ATTORNEY/AGENT INFORMATION:

; NAME: Mistrock, S. Leslie

; REGISTRATION NUMBER: 18,872

; REFERENCE/DOCKET NUMBER: 1101-174

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 790-9090

; TELEFAX: (212) 869-8864/9741

; INFORMATION FOR SEQ ID NO: 37:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 1400 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: DNA

; US-08-630-915A-37

Query Match 1.6%; Score 8; DB 4; Length 1400;

Best Local Similarity 100.0%; Pred. No. 1.3e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 AAAAGAA 301
|||||
Db 307 AAAAGAA 314

RESULT 28
US-09-548-372D-13

; Sequence 13, Application US/09548372D

; Patent No. 6420534

; GENERAL INFORMATION:

; APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 29915/62801

; CURRENT APPLICATION NUMBER: US/09/548,372D

; CURRENT FILING DATE: 2000-04-12

; PRIOR APPLICATION NUMBER: US 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 13

; LENGTH: 2088

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-548-372D-13

Query Match 1.6%; Score 8; DB 4; Length 2088;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 297 AGAAGAA 304
|||||
Db 692 AGAAGAA 699

RESULT 29
US-09-548-367D-13

; Sequence 13, Application US/09548367D

; Patent No. 6440698

; GENERAL INFORMATION:

; APPLICANT: GURNEY ET AL.

; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USE

; TITLE OF INVENTION: THEREOF

; FILE REFERENCE: 29915/6280H

; CURRENT APPLICATION NUMBER: US/09/548,367D

; CURRENT FILING DATE: 2000-04-12

; PRIOR APPLICATION NUMBER: US 60/155,493

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 09/404,133

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: PCT/US99/20881

; PRIOR FILING DATE: 1999-09-23

; PRIOR APPLICATION NUMBER: US 60/101,594

; PRIOR FILING DATE: 1998-09-24

; NUMBER OF SEQ ID NOS: 73

; SOFTWARE: Patentin version 3.1

; SEQ ID NO 13

; LENGTH: 2088

; TYPE: PRT

; ORGANISM: Homo sapiens

; US-09-548-367D-13

Query Match 1.6%; Score 8; DB 4; Length 2088;

Best Local Similarity 100.0%; Pred. No. 1.9e+02;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 297 AGAAGAA 304
|||||
Db 692 AGAAGAA 699

RESULT 30
US-09-738-884-1

; Sequence 1, Application US/09738884

; Patent No. 6391606

; GENERAL INFORMATION:

; APPLICANT: GUEGLER, Karl et al

; TITLE OF INVENTION: ISOLATED HUMAN PHOSPHOLIPASE PROTEINS,

; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN PHOSPHOLIPASE

; TITLE OF INVENTION: PROTEINS, AND USES THEREOF

; FILE REFERENCE: C1000849

; CURRENT APPLICATION NUMBER: US/09/738,884

```
; CURRENT FILING DATE: 2000-12-18
; NUMBER OF SEQ ID NOS: 5
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 2211
; TYPE: PRT
; ORGANISM: Human
US-09-738-884-1
```

```
Query Match          1.6%; Score 8; DB 4; Length 2211;
Best Local Similarity 100.0%; Pred. No. 2e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 297 AGGAGAA 304
Db 1361 AGGAGAA 1368
```

```
RESULT 31
US-09-627-650B-7
; Sequence 7, Application US/09627650B
; Patent No. 6406872
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: 21101.000903
; CURRENT APPLICATION NUMBER: US/09/627,650B
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/436,063
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107,727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2508
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-627-650B-7
```

```
Query Match          1.6%; Score 8; DB 4; Length 2508;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 299 GAAGAAA 306
Db 1184 GAAGAAA 1191
```

```
RESULT 32
US-09-436-063C-7
; Sequence 7, Application US/09436063C
; Patent No. 6407210
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: P-1095corrected
; CURRENT APPLICATION NUMBER: US/09/436,063C
; CURRENT FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 7
; LENGTH: 2508
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-436-063C-7
```

```
Query Match          1.6%; Score 8; DB 4; Length 2508;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 299 GAAGAAA 306
Db 1184 GAAGAAA 1191
```

```
RESULT 33
US-09-627-650B-3
; Sequence 3, Application US/09627650B
; Patent No. 6406872
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: 21101.000903
; CURRENT APPLICATION NUMBER: US/09/627,650B
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/436,063
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107,727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2544
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-627-650B-3
```

```
Query Match          1.6%; Score 8; DB 4; Length 2544;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 299 GAAGAAA 306
Db 1184 GAAGAAA 1191
```

```
RESULT 34
US-09-436-063C-3
; Sequence 3, Application US/09436063C
; Patent No. 6407210
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: P-1095corrected
; CURRENT APPLICATION NUMBER: US/09/436,063C
; CURRENT FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2544
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-436-063C-3
```

```
Query Match          1.6%; Score 8; DB 4; Length 2544;
Best Local Similarity 100.0%; Pred. No. 2.3e+02;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 299 GAAGAAA 306
Db 1184 GAAGAAA 1191
```

```
RESULT 35
US-09-627-650B-9
; Sequence 9, Application US/09627650B
; Patent No. 6406872
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: 21101.000903
; CURRENT APPLICATION NUMBER: US/09/627,650B
; CURRENT FILING DATE: 2000-07-28
; PRIOR APPLICATION NUMBER: 09/436,063
; PRIOR FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107,727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 2601
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-627-650B-9

Query Match
Best Local Similarity 1.6%; Score 8; DB 4; Length 2601;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 GAAGAAA 306
Db 1184 GAAGAAA 1191

RESULT 36
US-09-436-063C-9
; Sequence 9, Application US/09436063C
; Patent No. 6407210
; GENERAL INFORMATION:
; APPLICANT: Bamber, Bruce
; APPLICANT: Jorgensen, Erik
; TITLE OF INVENTION: Nematode Neuromuscular Junction GABA Receptors and
; FILE REFERENCE: P-1095corrected
; CURRENT APPLICATION NUMBER: US/09/436,063C
; CURRENT FILING DATE: 1999-11-08
; PRIOR APPLICATION NUMBER: 60/107727
; PRIOR FILING DATE: 1998-11-09
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 9
; LENGTH: 2601
; TYPE: PRT
; ORGANISM: Caenorhabditis elegans
US-09-436-063C-9

Query Match
Best Local Similarity 1.6%; Score 8; DB 4; Length 2601;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 GAAGAAA 306
Db 1184 GAAGAAA 1191

RESULT 37
US-09-336-447A-76
; Sequence 76, Application US/09336447A
; Patent No. 6310190
; GENERAL INFORMATION:
; APPLICANT: HANSEN, ERIC J.
; APPLICANT: AEBI, CHRISTOPH
; APPLICANT: COPE, LESLIE D.
; APPLICANT: MACIVER, ISOBEL
```

```
; APPLICANT: FISKE, MICHAEL J.
; APPLICANT: FREDENBURG, ROSS A.
; TITLE OF INVENTION: USP41 AND USP42 ANTIGENS OF MORAXELLA CATARRHALIS
; FILE REFERENCE: AMCY:024
; CURRENT APPLICATION NUMBER: US/09/336,447A
; CURRENT FILING DATE: 1999-06-21
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 76
; LENGTH: 3788
; TYPE: PRT
; ORGANISM: Moraxella catarrhalis
; FEATURE:
; NAME/KEY: MOD.RES
; LOCATION: (1036)..(3786)
; OTHER INFORMATION: Xaa = any
US-09-336-447A-76

Query Match
Best Local Similarity 1.6%; Score 8; DB 4; Length 3788;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 GAAGAAA 306
Db 3603 GAAGAAA 3610

RESULT 38
5178861-16
; Patent No. 5178861
; APPLICANT: VERGARA, ULISES;RUIZ, ANDRES;FERREIRA, ARTURO;
; NUSSENZWEIG, RUTH S.;NUSSENZWEIG, VICTOR N.
; TITLE OF INVENTION: CROSS-REACTIVE AND PROTECTIVE EPITOPES
; OF CIRCUMPOROZOITE PROTEINS
; NUMBER OF SEQUENCES: 18
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/370,241
; FILING DATE: 22-JUN-1989
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 115,634
; FILING DATE: 26-OCT-1987
; APPLICATION NUMBER: 649,903
; FILING DATE: 12-SEP-1984
; SEQ ID NO:16
; LENGTH: 11
5178861-16

Query Match
Best Local Similarity 1.4%; Score 7; DB 6; Length 11;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGG 299
Db 2 GAAAGG 8

RESULT 39
US-08-425-069-5
; Sequence 5, Application US/08425069
; Patent No. 5728810
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V.
; APPLICANT: Xu, Ming
; APPLICANT: Himan, Michael B.
; TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
; TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
; CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 No. 5728810th Washington Street
; CITY: Falls Church
; STATE: Virginia
```

COUNTRY: U.S.A.
ZIP: 22046
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/425,069
FILING DATE: 19-APR-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-106P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 205-8000
TELEFAX: (703) 205-8050
TELEX:
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Nephila clavipes
FEATURE:
NAME/KEY: Duplication
LOCATION: 1..21
OTHER INFORMATION: /label= repeat_unit
OTHER INFORMATION: /note= "spider silk protein repeat unit"
FEATURE:
NAME/KEY: Region
LOCATION: 1..6
OTHER INFORMATION: /label= alanine_stretch
OTHER INFORMATION: /note= "this segment of alanines in the repeat
OTHER INFORMATION: unit can also contain 7 alanine residues."
US-08-425-069-5
Query Match 1.4%; Score 7; DB 1; Length 21;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 294 AAAAGGA 300
Db 3 AAAAGGA 9
RESULT 40
US-08-317-844B-5
Sequence 5, Application US/08317844B
Patent No. 5989894
GENERAL INFORMATION:
APPLICANT: Lewis, Randolph V.
APPLICANT: Xu, Ming
APPLICANT: Hinman, Michael B.
TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
TITLE OF INVENTION: CONTAINING THE ISOLATED DNA, AND PRODUCTS THEREOF
NUMBER OF SEQUENCES: 62
CORRESPONDENCE ADDRESS:
ADDRESSEE: Birch, Stewart, Kolasch & Birch
STREET: 301 No. 5989894th Washington Street
CITY: Falls Church
STATE: Virginia
COUNTRY: U.S.A.
ZIP: 22046
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/317,844B
FILING DATE: 04-OCT-1994
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Murphy Jr., Gerald M
REGISTRATION NUMBER: 28,977
REFERENCE/DOCKET NUMBER: 1447-105P
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 241-1300
TELEFAX: (703) 241-2848
TELEX: 248345
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 21 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FRAGMENT TYPE: internal
ORIGINAL SOURCE:
ORGANISM: Nephila clavipes
FEATURE:
NAME/KEY: Duplication
LOCATION: 1..21
OTHER INFORMATION: /label= repeat_unit
OTHER INFORMATION: /note= "spider silk protein repeat unit"
FEATURE:
NAME/KEY: Region
LOCATION: 1..6
OTHER INFORMATION: /label= alanine_stretch
OTHER INFORMATION: /note= "this segment of alanines in the repeat
OTHER INFORMATION: unit can also contain 7 alanine residues."
US-08-317-844B-5
Query Match 1.4%; Score 7; DB 2; Length 21;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 294 AAAAGGA 300
Db 3 AAAAGGA 9
RESULT 41
US-07-987-286-3
Sequence 3, Application 07/987286
Patent No. 5500366
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, GREGORY J
APPLICANT: GECEZY, ANDREW F
TITLE OF INVENTION: T-CELL EPITOPES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K STREET, N.W.
CITY: WASHINGTON, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 07/987,286
FILING DATE: 16-MAR-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU91/00429
FILING DATE: 17-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PK2361

FILING DATE: 18-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: BENT, STEPHEN A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/175/CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-07-987-286-3

Query Match 1.4%; Score 7; DB 1; Length 23;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 289 GAAVGAA 295
Db 10 GAAVGAA 16

RESULT 42
US-08-614-626-3
Sequence 3, Application US/08614626
Patent No. 5928644
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, GREGORY J
APPLICANT: GECZY, ANDREW F
TITLE OF INVENTION: T-CELL EPITOPES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K STREET, N.W.
CITY: WASHINGTON, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/614,626
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/987,286
FILING DATE: 16-MAR-1993
APPLICATION NUMBER: PCT/AU91/00429
FILING DATE: 17-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PK2361
FILING DATE: 18-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: BENT, STEPHEN A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/175/CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 23 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-614-626-3

Query Match 1.4%; Score 7; DB 2; Length 23;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 289 GAAVGAA 295
Db 10 GAAVGAA 16

RESULT 43
US-07-987-286-11
Sequence 11, Application 07/987286
Patent No. 5500366
GENERAL INFORMATION:
APPLICANT: RUSSELL-JONES, GREGORY J
APPLICANT: GECZY, ANDREW F
TITLE OF INVENTION: T-CELL EPITOPES
NUMBER OF SEQUENCES: 26
CORRESPONDENCE ADDRESS:
ADDRESSEE: FOLEY & LARDNER
STREET: 3000 K STREET, N.W.
CITY: WASHINGTON, D.C.
COUNTRY: USA
ZIP: 20007-5109
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: 07/987,286
FILING DATE: 16-MAR-1993
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/AU91/00429
FILING DATE: 17-SEP-1991
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PK2361
FILING DATE: 18-SEP-1990
ATTORNEY/AGENT INFORMATION:
NAME: BENT, STEPHEN A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 16786/175/CHAC
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202 672 5300
TELEFAX: 202 672 5399
TELEX: 904136
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 25 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
FEATURE:
NAME/KEY: Modified-site
LOCATION: 1
OTHER INFORMATION: /note= "RESIDUE 1 IS PYROGLUTAMIC
OTHER INFORMATION: ACID"
FEATURE:
NAME/KEY: Modified-site
LOCATION: 25
OTHER INFORMATION: /note= "RESIDUE 25 IS CYSTEINAMIDE"
US-07-987-286-11

Query Match 1.4%; Score 7; DB 1; Length 25;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 289 GAAVGAA 295
Db 11 GAAVGAA 17

RESULT 44
US-07-987-286-24
; Sequence 24, Application 07/987286
; Patent No. 5500366
; GENERAL INFORMATION:
; APPLICANT: RUSSELL-JONES, GREGORY J
; APPLICANT: GECZY, ANDREW F
; TITLE OF INVENTION: T-CELL EPITOPES
; NUMBER OF SEQUENCES: 26
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FOLEY & LARDNER
; STREET: 3000 K STREET, N.W.
; CITY: WASHINGTON, D.C.
; COUNTRY: USA
; ZIP: 20007-5109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: 07/987,286
; FILING DATE: 16-MAR-1993
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/AU91/00429
; FILING DATE: 17-SEP-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PR2361
; FILING DATE: 18-SEP-1990
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, STEPHEN A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 16786/175/CHAC
; TELEPHONE: 202 672 5300
; TELEFAX: 202 672 5399
; TELEX: 904136
; INFORMATION FOR SEQ. ID NO: 24:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-07-987-286-24

Query Match 1.4%; Score 7; DB 1; Length 25;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 289 GAAVGA 295
| | | | |
DB 12 GAAVGA 18

RESULT 45
US-08-425-069-38
; Sequence 38, Application US/08425069
; Patent No. 5728810
; GENERAL INFORMATION:
; APPLICANT: Lewis, Randolph V.
; APPLICANT: Xu, Ming
; APPLICANT: Hinman, Michael B.
; TITLE OF INVENTION: ISOLATED DNA CODING FOR SPIDER SILK
; TITLE OF INVENTION: PROTEIN, A REPLICABLE VECTOR AND A TRANSFORMED CELL
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Birch, Stewart, Kolasch & Birch
; STREET: 301 No. 5728810th Washington Street
; CITY: Falls Church
; STATE: Virginia

COUNTRY: U.S.A.
ZIP: 22046
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/425,069
; FILING DATE: 19-APR-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Murphy Jr., Gerald M
; REGISTRATION NUMBER: 28,977
; REFERENCE/DOCKET NUMBER: 1447-106P
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 205-8000
; TELEFAX: (703) 205-8050
; TELEX:
; INFORMATION FOR SEQ. ID NO: 38:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 25 amino acids
; TYPE: amino acid
; STRANDEDNESS: not relevant
; TOPOLOGY: not relevant
; MOLECULE TYPE: peptide
; HYPOTHEICAL: NO
; FRAGMENT TYPE: Internal
; ORIGINAL SOURCE:
; ORGANISM: nephila clavipes
; FEATURE:
; NAME/KEY: peptide
; LOCATION: 1..25
; OTHER INFORMATION: /label=silk1_repeat
; US-08-425-069-38

Query Match 1.4%; Score 7; DB 1; Length 25;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 AAAAGA 300
| | | | |
DB 7 AAAAGA 13

Search completed: January 27, 2003, 16:37:39
Job time : 41 secs

FT /note= "Mature ORP-3"
 XX WO200144448-A2.
 XX 21-JUN-2001.
 PD
 XX 07-DEC-2000; 2000WO-US33158.
 XX 16-DEC-1999; 99US-0172367.
 PR
 XX (INCY-) INCYTE GENOMICS INC.
 PA
 XX Yue H, Lal P, Tang YT, Hillman JL, Baughn MR, Azimzal Y, Lu DM;
 PI
 XX WPI: 2001-390245/41.
 DR
 XX N-PSDB; AAH24225.
 DR
 PT Novel human oxidoreductase protein (ORP) useful for diagnosing,
 PT treating and preventing cell proliferative, neurological, viral,
 PT reproductive and autoimmune/inflammatory disorders associated with
 PT abnormal expression of ORP -
 XX
 XX Claim 1: Page 99-100; 136pp; English.
 PS
 XX Sequences AAB73668-AAB73694 represent 27 novel human oxidoreductase
 CC proteins, designated ORP-1 to ORP-27 respectively, and sequences
 CC AAH24223-AAH24249 represent cDNAs encoding ORP-1 to ORP-27. Human ORP
 CC proteins and nucleic acids are useful for diagnosing, treating or
 CC preventing cell proliferative disorders (e.g. arteriosclerosis,
 CC cirrhosis, psoriasis, cancers); endocrine disorders (e.g., type I or II
 CC diabetes mellitus, diabetes insipidus, dwarfism, hirsutism, amenorrhoea,
 CC osteoporosis); metabolic disorders (e.g., obesity, phenylketonuria,
 CC hypercholesterolaemia); reproductive disorders (e.g., infertility,
 CC ovulatory and menstrual cycle defects, endometriosis, polycystic
 CC ovary disease, disruption of spermatogenesis, impotence); neurological
 CC disorders (e.g., epilepsy, stroke, Alzheimer's disease, Huntington's
 CC disease, Parkinson's disease, meningitis, Creutzfeldt-Jakob disease,
 CC cerebral palsy, muscular dystrophy, mood, anxiety and schizophrenic
 CC disorders); viral, bacterial, fungal and parasitic infections; and
 CC autoimmune/inflammatory disorders such as acquired immunodeficiency
 CC syndrome (AIDS), allergies, asthma, Crohn's disease, atopic dermatitis,
 CC gout, multiple sclerosis, rheumatoid arthritis or ulcerative colitis.
 CC Human ORP proteins and nucleotides can be used to identify compounds
 CC which modulate their activity or expression. ORP nucleic acid sequences
 CC may also be used for assessing the toxicity of a test compound, to detect
 CC upstream sequences such as promoters and regulatory elements, and to
 CC create knock out or knock in animals or transgenic animals to model
 CC human disease. Oligonucleotide primers derived from ORP gene sequences
 CC may be used to detect single nucleotide polymorphisms (SNPs) and for
 CC mapping the naturally occurring genomic sequences. Antibodies specific
 CC for ORP proteins may be used in the diagnosis of disorders associated
 CC with aberrant ORP expression, in assays to monitor patients being treated
 CC with ORP or modulators thereof, and for assessing toxicity of potential
 CC drugs.
 CC
 XX
 XX Sequence 555 AA:
 SQ
 Query Match 1.8%; Score 9; DB 22; Length 555;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 280 GAGIAGIAA 288
 |||||||
 DB 31 GAGIAGIAA 39
 RESULT 45
 ABB58615
 ID ABB58615 standard; Protein; 696 AA.
 XX
 XX ABB58615;
 AC
 XX
 DT 26-MAR-2002 (first entry)

XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 2637.
 XX
 XX Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 KW
 XX Drosophila melanogaster.
 OS
 XX WO200171042-A2.
 PN
 XX 27-SEP-2001.
 PD
 XX 23-MAR-2001; 2001WO-US09231.
 XX
 XX PF
 PR 23-MAR-2000; 2000US-191637P.
 PR 11-JUL-2000; 2000US-0614150.
 XX
 XX (PEKE) PE CORP NY.
 PA
 XX Venter JC, Adams M, Li PWD, Myers EW;
 PI
 XX WPI: 2001-656860/75.
 DR
 XX N-PSDB; ABL02718.
 DR
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions -
 XX
 PS Disclosure; SEQ ID NO 2637; 21pp + Sequence Listing; English.
 XX
 XX The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL01840-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (AB57737-AB572072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pct_sequences.
 CC
 XX
 XX Sequence 696 AA:
 SQ
 Query Match 1.8%; Score 9; DB 22; Length 696;
 Best Local Similarity 100.0%; Pred. No. 31;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 293 GAAAGCAA 301
 |||||||
 DB 88 GAAAGCAA 96

Search completed: January 27, 2003, 16:34:50
 Job time : 60 secs

PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Dmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R.
 DR WPI: 2001-476283/51.
 DR N-PSDB; AAK52679.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX
 PS Claim 20; Page 280; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAW78323-AAW80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activating
 CC activity/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAW80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX
 SQ Sequence 518 AA;
 XX
 Query Match 1.8%; Score 9; DB 22; Length 518;
 Best Local Similarity 100.0%; Pred. No. 24;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 280 GAGLAGLAA 288
 IIIIIIIIII
 Db 37 GAGLAGLAA 45
 XX
 RESULT 43
 AAB12164
 ID AAB12164 standard; Protein; 555 AA.
 XX
 AC AAB12164;
 XX
 DT 02-FEB-2001 (first entry)
 XX
 DE Hydrophobic domain protein from clone HP10673 isolated from thymus cells.
 XX
 KW Human; secreted protein; membrane protein; hydrophobic domain;
 KW proliferation control; differentiation induction; material transport;
 KW biophylaxis; signal receptor; ion channel; transporter; immunostimulant;
 KW immunosuppressant; haematopoiesis regulator; chemotactic; chemokinetic;
 KW hemostatic; thrombolytic; anti-inflammatory; tumour inhibition;
 KW autoimmune disease; Alzheimer's disease; Parkinson's disease; cancer.
 XX
 OS Homo sapiens.
 XX
 PN WO200029448-A2.
 XX
 PD 25-MAY-2000.
 XX
 PF 17-NOV-1999; 99WO-JP06412.
 XX
 PR 17-NOV-1998; 98JP-0326255.
 PR 22-DEC-1998; 98JP-0364315.
 PR 16-MAR-1999; 99JP-0069811.
 PR 27-APR-1999; 99JP-0119299.
 PR 19-MAY-1999; 99JP-0138169.

XX
 PA (SAGA) SAGAMI CHEM RES CENT.
 PA (PROT-) PROTEGENE INC.
 XX
 PI Kato S, Kimura T;
 PI
 DR WPI: 2000-387753/33.
 DR N-PSDB; AAA62067, AAA62077.
 XX
 PT Proteins comprising hydrophobic regions, such as secretory and membrane
 PT proteins, useful in research and diagnostics and having various
 PT activities e.g. immunomodulatory, antiinflammatory, chemokinetic,
 PT hemostatic, thrombolytic -
 XX
 PS Claim 1; Page 376-378; 410pp; English.
 XX
 CC Secretory proteins play important roles in the proliferation control, the
 CC differentiation induction, the material transport and the biophylaxis of
 CC cells. Membrane proteins have important roles as signal receptors, ion
 CC channels and transporters. The present sequence is a human protein which
 CC has at least one hydrophobic domain. This protein may be a secretory or a
 CC membrane protein. The present protein may have cytokine and cell
 CC proliferation/differentiation activity, immune stimulating or suppressing
 CC activity, haematopoiesis activity, tissue growth activity,
 CC activity/inhibin activity, chemotactic/chemokinetic activity, haemostatic
 CC inhibition activity, anti-inflammatory activity and tumour
 CC treatment activity. The present protein could therefore be used for
 CC disease, and cancer.
 XX
 SQ Sequence 555 AA;
 XX
 Query Match 1.8%; Score 9; DB 21; Length 555;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 280 GAGLAGLAA 288
 IIIIIIIIII
 Db 31 GAGLAGLAA 39
 XX
 RESULT 44
 AAB73670
 ID AAB73670 standard; Protein; 555 AA.
 XX
 AC AAB73670;
 XX
 DT 11-SEP-2001 (first entry)
 XX
 DE Human oxidoreductase protein ORP-3.
 XX
 KW Human oxidoreductase protein; ORP; cell proliferative disorder;
 KW arteriosclerosis; cirrhosis; psoriasis; cancer; endocrine disorder;
 KW diabetes mellitus; diabetes insipidus; dwarfism; hirsutism; amenorrhoea;
 KW osteoporosis; metabolic disorder; obesity; phenylketonuria;
 KW hypercholesterolaemia; reproductive disorder; infertility;
 KW ovulatory defect; menstrual cycle defect; endometriosis;
 KW polycystic ovary disease; spermatogenesis disruption; impotence;
 KW neurological disorder; epilepsy; stroke; Alzheimer's disease;
 KW Huntington's disease; Parkinson's disease; Creutzfeldt-Jakob disease;
 KW meningitis; cerebral palsy; muscular dystrophy; mood disorder; anxiety;
 KW schizophrenic disorder; infection; autoimmune disorder;
 KW inflammatory disorder; acquired immunodeficiency syndrome; AIDS; asthma;
 KW allergy; Crohn's disease; atopic dermatitis; gout; multiple sclerosis;
 KW rheumatoid arthritis; ulcerative colitis; drug screening;
 KW toxicity screening; transgenic animal; SNP detection; gene therapy.
 XX
 OS Homo sapiens.
 XX
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..40
 FT /label= signal_peptide
 FT Protein 41..555

OS Chlamydia trachomatis.
 XX
 PN MO9928475-A2.
 XX
 PD 10-JUN-1999.
 XX
 PF 27-NOV-1998; 98WO-IB01939.
 XX
 PR 04-NOV-1998; 98US-0107077.
 PR 28-NOV-1997; 97FR-0015041.
 PR 17-DEC-1997; 97FR-0016034.
 XX
 PA (GEST) GENSET.
 XX
 PI Griffais R;
 XX
 DR WPI; 1999-371125/31.
 XX
 PT Genome sequence of Chlamydia trachomatis
 XX
 PS Disclosure; Page 1271-1272; 1755pp; English.
 XX
 CC AAY36754-y37949 are encoded by open reading frames (ORFs) of the genome
 CC of Chlamydia trachomatis (see AA201425). The polypeptides can be used as
 CC vaccines against Chlamydia trachomatis. Antisense and ribozyme sequences
 CC can also be used to control growth of the microorganism. Chlamydia
 CC trachomatis is responsible for a large number of diseases, e.g. eye
 CC diseases such as conventional trachoma, nongonococcal trachoma,
 CC paratrachoma, and inclusion conjunctivitis; genital diseases such as
 CC nongonococcal urethritis, epididymitis, cervicitis, salpingitis,
 CC peritonitis, Bartholinitis; pneumonia in breast feeding infants;
 CC and venereal lymphogranulomatosis. The polypeptides of the invention
 CC may be of use in treating these diseases.
 XX
 SQ Sequence 481 AA;
 Query Match 1.8%; Score 9; DB 20; Length 481;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 447 FMOQASKIA 455
 |||||||||
 DB 446 FMOQASKIA 454
 RESULT 41
 AAM78562
 ID AAM78562 standard; Protein; 513 AA.
 XX
 AC AAM78562;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 1224.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN MO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US04098.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.

PR 15-SEP-2000; 2000US-0663561.
 PR 20-OCT-2000; 2000US-0693325.
 PR 30-NOV-2000; 2000US-0728422.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y, Ma Y;
 PI Zhao QA, Wang D, Zhang J, Ren F, Chen R, Wang ZW;
 PI Xue AJ, Yang Y, Wejhrman T, Goodrich R.
 XX
 DR WPI; 2001-476283/51.
 DR N-PSDB; AAK51695.
 XX
 PT Nucleic acids encoding polypeptides with cytokine-like activities,
 PT useful in diagnosis and gene therapy -
 XX
 PS Claim 20; Page 3481-3482; 6221pp; English.
 XX
 CC The invention relates to polynucleotides (AAK51456-AAK53435) and the
 CC encoded polypeptides (AAM78323-AAK80302) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activity/inhibn activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation.
 CC Note: Records for SEQ ID NO 2110 (AAK52581), 2111 (AAK52582) and 3666
 CC (AAM80020) are omitted as the relevant pages from the sequence listing
 CC were missing at the time of publication.
 XX
 SQ Sequence 513 AA;
 Query Match 1.8%; Score 9; DB 22; Length 513;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 280 GAGIAGLAA 288
 |||||||||
 DB 31 GAGIAGLAA 39
 RESULT 42
 AAM79546
 ID AAM79546 standard; Protein; 518 AA.
 XX
 AC AAM79546;
 XX
 DT 06-NOV-2001 (first entry)
 XX
 DE Human protein SEQ ID NO 3192.
 XX
 KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 KW tissue growth factor; immunomodulatory; cancer; leukaemia;
 KW nervous system disorder; arthritis; inflammation.
 XX
 OS Homo sapiens.
 XX
 PN MO200157190-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US04098.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 PR 20-JUN-2000; 2000US-0598075.
 PR 19-JUL-2000; 2000US-0620325.
 PR 01-SEP-2000; 2000US-0654936.
 PR 15-SEP-2000; 2000US-0663561.

```

PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157111.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0158923.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161970.
PR 28-OCT-1999; 99US-0161982.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

Query Match 1.8%; Score 9; DB 21; Length 382;
 Best Local Similarity 100.0%; Pred No. 18;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 280 GAGIAGLAA 288
 |||||||
 DB 56 GAGIAGLAA 64

```

RESULT 39
ABR92256
ID ABR92256 standard; Protein; 477 AA.
XX
AC ABR92256;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 1467.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP09892.
XX
PR 28-AUG-2001; 2001WO-EP09892.
XX
PA (FARB ) BAYER AG.
XX
PI Tietjen K, Weidner M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms -
XX
PS Claim 5; SEQ ID NO 1467; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins
CC (ABR9790-ABR94016) for herbicidally active compounds, comprising
CC aligning and comparing nucleic acid or amino acid sequences from plant
CC with nucleic acid or amino acid sequences from non-plant organisms using
CC suitable search parameters, where plant sequences having an E-value
CC greater by a factor of 3 than the E-value of most similar non-plant
CC sequences are selected. The polypeptides or nucleic acids encoding them
CC are useful for identifying modulators. The identified modulators are
CC useful as herbicides.
XX
SQ Sequence 477 AA;
XX
Query Match 1.8%; Score 9; DB 23; Length 477;
Best Local Similarity 100.0%; Pred No. 22;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 280 GAGIAGLAA 288
|||||||
DB 53 GAGIAGLAA 61
RESULT 40
AAV37632
ID AAV37632 standard; Protein; 481 AA.
XX
AC AAV37632;
XX
DT 07-OCT-1999 (first entry)
XX
DE Protein which is specific to Chlamydia trachomatis.
XX
KW Vaccine; eye disease; conventional trachoma; nongonococcal trachoma;
KW paratrachoma; inclusion conjunctivitis; genital disease; peritrichitis;
KW nongonococcal urethritis; epididymitis; cervicitis; salpingitis;
KW Bartholinitis; pneumonia; venereal lymphogranulomatosis.
XX

```

Query Match 1.8%; Score 9; DB 21; Length 379;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 GAGLAGLAA 288
|||||||
DB 53 GAGLAGLAA 61

RESULT 38
AAG10363
ID AAG10363 standard; Protein; 382 AA.
XX
AC AAG10363;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 8653.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
hybridisation assay; genetic mapping; gene expression control; promoter;
termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP103405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129645.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139617.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142820.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.

PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144329.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144652.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145086.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 10-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 26-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

CC haemolytic anaemia, autoimmune thyroiditis, diabetes mellitus,
CC Crohn's disease, multiple sclerosis, rheumatoid arthritis and ulcerative
CC colitis; (c) cardiovascular disorders such as myocardial ischaemias;
CC (d) wound healing; (e) neurological diseases e.g. cerebral anoxia and
CC epilepsy; and (f) infectious diseases such as viral, bacterial, fungal
CC and parasitic infections.

XX
SQ Sequence 172 AA;

Query Match 1.8%; Score 9; DB 21; Length 172;
Best Local Similarity 100.0%; Pred. No. 9;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 296 AAGGAGAA 304
1111111111

DB 42 AAGGAGAA 50

RESULT 36
ABG05860

ID ABG05860 standard; Protein; 354 AA.

AC ABG05860;

DT 13-FEB-2002 (first entry)

DE Novel human diagnostic protein #5851.

XX Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder.

OS Homo sapiens.

XX WO200175067-A2.

PN 11-OCT-2001.

PD 30-MAR-2001; 2001WO-US08631.

PF 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

XX (HYSE-) HYSEQ INC.

XX Drmanac RT, Liu C, Tang YT;

XX WPI: 2001-639362/73.

DR N-PSDB: AAS70047.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity -

PS Claim 20; SEQ ID No 36219; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant production of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful in medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. ABG0010-ABG30377 represent novel human

CC diagnostic amino acid sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 354 AA;

Query Match 1.8%; Score 9; DB 22; Length 354;
Best Local Similarity 100.0%; Pred. No. 17;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 286 LAAGAAVGA 294
1111111111

DB 149 LAAGAAVGA 157

RESULT 37
AAG10364

ID AAG10364 standard; Protein; 379 AA.

AC AAG10364;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 8654.

XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.

XX Arabidopsis thaliana.

OS EP1033405-A2.

XX PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

PR 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132487.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135553.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

CC kidney failure) and to delay or prevent the development of PKD-associated
 CC hypertension, bleeding into the cysts, pain, and renal insufficiency
 CC associated with the progression of cystic disease. In an exemplification
 CC of the invention, the expression profile of representative genes was
 CC examined in a rat model of polycystic kidney disease and compared with
 CC gene expression in normal rats. The PRBR gene (also referred to as clone
 CC PKD42) was found to be overexpressed by a factor of 2 in the rat model of
 CC PKD, while the DBI gene was underexpressed by a factor of 1.9. The
 CC present sequence represents the amino acid sequence of human PRBR.
 CC Note: The present sequence is not given in the specification although it
 CC is referred to as SEQ ID NO:3 and is stated as being given in figure 2.
 CC This sequence was obtained using the DNA sequence given in figure 2.

XX Sequence 169 AA;

Query Match 1.8%; Score 9; DB 21; Length 169;

Best Local Similarity 100.0%; Pred. No. 8.9;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 GAAATTTVA 310

Db 117 GAAATTTVA 125

RESULT 34

AA93604
 ID AAY93604 standard; Protein: 169 AA.

XX AC AAY93604;

XX DT 25-SEP-2000 (first entry)

XX DE Protein encoded by a human peripheral benzodiazepine receptor gene.

XX KW Differentially expressed human gene: cardiac disease; kidney disease;
 KW inflammatory disease; I-80; prostacyclin-stimulating factor; isf-2;
 KW tissue specific mRNA; insulin-like growth factor binding protein 6;
 KW OSF-1; gas-1; YMP; BTG2; pre-B cell stimulating factor homologue; SDF1a;
 KW peripheral benzodiazepine receptor; annexin II cellular ligand; p11;
 KW congenital heart failure; dilated congestive cardiomyopathy;
 KW hypertrophic cardiomyopathy; restrictive cardiomyopathy;
 KW mitral valve disease; aortic valve disease; tricuspid valve disease;
 KW angina pectoris; myocardial infarction; cardiac arrhythmia;
 KW pulmonary hypertension; arterial hypertension; renovascular hypertension;
 KW arteriosclerosis; atherosclerosis; cardiac tumour.

XX OS Homo sapiens.

XX PN WO200035473-A2.

XX PD 22-JUN-2000.

XX PF 15-DEC-1999; 99WO-US29941.

XX PR 18-DEC-1998; 98US-0113008.

XX PA (SCIO-) SCIOS INC.

XX PI Stanton LW, White RT, Damm DL, Lewicki JA, Joly A, Schreiner GF,

XX DR WPI; 2000-451904/39.

XX DR N-PSDB; AAA46678.

XX PT Preventing, diagnosing and treating cardiac, kidney and inflammatory
 PT disorders using cardiac genes that are differentially expressed in
 PT disease states such as cardiac arrhythmia and arteriosclerosis -

XX PS Disclosure: Fig 8K; 170pp; English.

XX CC AAY93594-Y93605 are encoded by differentially expressed human genes,
 CC associated with disease states and disorders. The specification
 CC describes methods preventing, diagnosing and treating cardiac, kidney
 CC and inflammatory diseases associated with inappropriate expression of

CC differentially expressed cardiac, kidney and inflammatory genes
 CC (e.g. AAA46668-79). These genes include I-80, prostacyclin-stimulating
 CC factor, isf-2, tissue specific mRNA, insulin-like growth factor
 CC binding protein 6, OSF-1, gas-1, YMP, BTG2, pre-B cell stimulating
 CC factor homologue (SDF1a), peripheral benzodiazepine receptor, and
 CC cellular ligand of annexin II (p11), respectively. These diseases
 CC include congenital heart failure, dilated congestive cardiomyopathy,
 CC hypertrophic cardiomyopathy, restrictive cardiomyopathy, mitral valve
 CC disease, aortic valve disease, tricuspid valve disease, angina pectoris,
 CC myocardial infarction, cardiac arrhythmia, pulmonary hypertension,
 CC arterial hypertension, renovascular hypertension, arteriosclerosis,
 CC atherosclerosis and/or cardiac tumours.

XX Sequence 169 AA;

Query Match 1.8%; Score 9; DB 21; Length 169;

Best Local Similarity 100.0%; Pred. No. 8.9;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 GAAATTTVA 310

Db 117 GAAATTTVA 125

RESULT 35

AAB38041
 ID AAB38041 standard; Peptide: 172 AA.

XX AC AAB38041;

XX DT 31-JAN-2001 (first entry)

XX DE Fragment of human secreted protein encoded by gene 10 clone HMHGFP71.

XX KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
 KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; anticancer;
 KW vulnerary; anticoagulant; antibacterial; antifungal; antiparasitic;
 KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
 KW neurological disease; infection; human; secreted protein.

XX OS Homo sapiens.

XX PN WO200055371-A1.

XX PD 21-SEP-2000.

XX PE 16-MAR-2000; 2000WO-US06783.

XX PR 18-MAR-1999; 99US-0125055.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ruben SM, Ni J, Ebner R, Rosen CA, Shi Y, Birse C, Florence K;

XX DR Komatsoulis G, Lafleur DM, Moore PA, Olsen HS, Young PE;

XX DR WPI; 2000-594448/56.

XX PT New nucleic acid molecules encoding 27 human secreted proteins for
 PT diagnosing, preventing, treating or ameliorating medical conditions and
 PT used as food additives or preservatives -

XX PS Disclosure: Page 27; 453pp; English.

XX CC Sequences AAB37984-B38019 represent the amino acid sequences of 27
 CC human secreted proteins encoded by the genes AAC69084-C69119. The genes
 CC and proteins are useful for preventing, ameliorating or treating medical
 CC conditions, e.g. by protein or gene therapy. The genes are isolated from
 CC a range of human tissues disclosed in the specification. The nucleic
 CC acids, proteins, antibodies and (ant)agonists are useful in the
 CC diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer, and other cancers of the adrenal gland, bone, bone
 CC marrow, breast, gastrointestinal tract, liver, lung, or urogenital;
 CC (b) immune disorders e.g. Addison's disease, allergies, autoimmune

PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 1.8%; Score 9; DB 21; Length 159;
 Best Local Similarity 100.0%; Pred. No. 8.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGAA 301
 Db 129 GAAAGGAA 137

RESULT 32
 AAY42438
 ID AAY42438 standard; Protein; 169 AA.
 AC AAY42438;

DT 08-DEC-1999 (first entry)
 XX
 DE Amino acid sequence of partial PBR from MDA-231 and MCF-7.

KW Cancer; cell proliferation; invasive tumor; aggressive tumor; cytostatic;
 KM antiproliferative; hypoproliferative; detection; diagnosis.

OS Homo sapiens.

FT Key Location/Qualifiers
 FT Region 1.26
 FT /note="nucleotide sequence unknown"

PN WO949316-A2.

PD 30-SEP-1999.

PF 25-MAR-1999; 99MO-US06515.

PR 25-MAR-1998; 98US-0047652.

PA (GEO) UNIV GEORGETOWN MEDICAL CENT.

PI Papadopoulos V, Culty M;

DR MPI: 1999-580494/49.

DR N-PSDB: AA222636, AA222637.

PT Agents that reduce peripheral-type benzodiazepine receptor function,
 PT useful for detecting and creating aggressive phenotype breast cancer
 XX
 PS Claim 32; Page 74-75; 76pp; English.

CC This is the amino acid sequence of the peripheral-type benzodiazepine
 CC receptor, derived from the human breast cancer cell lines, MCF-7 and
 CC MDA-231. The amino acid sequence differs from the wildtype PBR sequence
 CC with the replacement of histidine 162 with arginine and replacement
 CC of alanine 147 with a threonine.

CC This invention provides the means of peripheral-type benzodiazepine
 CC receptor (PBR) modulation.

CC Molecules that inhibit PBR function are used to treat diseases
 CC mediated by increased cell proliferation, particularly cancers (e.g.
 CC cancers of the lung, stomach, prostate, ovary, testes, pancreas, liver,
 CC cervix, vagina, skin, esophagus, nasopharynx, oropharynx, brain,
 CC central and peripheral nervous systems, bone and cartilage), especially

CC breast cancer. PBR, or a vector encoding PBR may be used to treat
 CC diseases mediated by reduced cell proliferation, such as developmental
 CC retardation.

XX Sequence 169 AA;

Query Match 1.8%; Score 9; DB 20; Length 169;
 Best Local Similarity 100.0%; Pred. No. 8.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 GAAATTVA 310
 Db 117 GAAATTVA 125

RESULT 33
 AAB11966
 ID AAB11966 standard; Protein; 169 AA.
 AC AAB11966;

DT 22-NOV-2000 (first entry)

DE Human peripheral-type benzodiazepine receptor (PTBR).

KW polycystic kidney disease; PKD; hypertension; renal failure; form;
 KM insufficiency; autosomal dominant form; autosomal recessive form;
 KM cystic disease; PTBR ligand; differential expression; antiproliferative;
 KM hypotensive.

OS Homo sapiens.

PN WO200037085-A1.

PD 29-JUN-2000.

PF 15-DEC-1999; 99MO-US29763.

PR 18-DEC-1998; 98US-0113008.

PR 26-MAY-1999; 99US-0136208.

PA (SCIO-) SCIOS INC.

PI (STAN/) STANTON L W.

DR MPI: 2000-442524/38.

DR N-PSDB: AAA72048.

PT Treatment and prevention of cystic disease, specifically polycystic
 PT kidney disease, by administering ligand of peripheral benzodiazepine
 PT receptor
 XX
 PS Example 1; Page -: 37pp; English.

CC The invention relates to a method for the treatment and prevention of
 CC diseases that involve cyst formation, particularly polycystic kidney
 CC disease (PKD), comprising administering a ligand of a peripheral-type
 CC benzodiazepine receptor (PTBR). The PTBR ligand can also be used for the
 CC treatment of hypertension associated with PKD. The invention also
 CC encompasses the diagnosis and/or prognosis of cystic disease, or the
 CC diagnosis of a predisposition to cystic disease, via the monitoring of
 CC changes in expression levels of PTBR or an endogenous ligand of PTBR,
 CC such as the diazepam binding inhibitor (DBI). In PKD, the PTBR gene is
 CC overexpressed, while the natural ligand is underexpressed. The PTBR
 CC ligands of the invention inhibit the proliferation of epithelial cells
 CC from PKD kidneys, i.e., they arrest cell growth but do not kill cells.
 CC The method is used to treat diseases associated with cyst formation,
 CC particularly both the autosomal dominant and recessive forms of PKD.
 CC The methods can also be used for the diagnosis and prognosis of a cystic
 CC disease or susceptibility to it. The methods may also be used to
 CC delay or prevent cyst expansion (and therefore subsequent progression to

PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138099.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 17-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-01452913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151338.
PR 01-SEP-1999; 99US-0151430.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.

DB 56 GAAAGCAA 64

RESULT 30

ABG41659

ID ABG41659 standard; Peptide: 153 AA.

XX ABG41659;

XX 19-AUG-2002 (first entry)

DE Human peptide encoded by genome-derived single exon probe SEQ ID 31324.

XX Human: single exon probe; asthma; lung cancer; COPD; IHD;

KW Chronic obstructive pulmonary disease; interstitial lung disease;

KW familial idiopathic pulmonary fibrosis; neurofibromatosis;

KW tuberous sclerosis; Gaucher's disease; Niemann-Pick disease;

KW Hermansky-Pudlak syndrome; sarcoidosis; pulmonary haemostasis;

KW pulmonary histiocytosis; lymphangioleiomyomatosis; Karagener syndrome;

KW pulmonary alveolar proteinosis; fibrocystic pulmonary dysplasia;

KW primary ciliary dyskinesia; pulmonary hypertension;

KW hyaline membrane disease.

XX Homo sapiens.

OS MO200186003-A2.

PN 15-NOV-2001.

XX 30-JAN-2001; 2001WO-US00665.

PF 04-FEB-2000; 2000US-180312P.

XX 26-MAY-2000; 2000US-207456P.

PR 30-JUN-2000; 2000US-0608408.

PR 03-AUG-2000; 2000US-0632366.

PR 21-SEP-2000; 2000US-234687P.

PR 27-SEP-2000; 2000US-236359P.

PR 04-OCT-2000; 2000GB-0024263.

XX (MOLE-) MOLECULAR DYNAMICS INC.

XX Penn SG, Hanzel DK, Chen W, Rank DR;

PI WPI; 2002-114183/15.

DR Spatially-addressable set of single exon nucleic acid probes, used to

PT measure gene expression in human lung samples -

XX Claim 27; SEQ ID No 31324; 634pp; English.

XX The invention relates to a spatially-addressable set of single exon

CC nucleic acid probes for measuring gene expression in a sample derived

CC from human lung comprising single exon nucleic acid probes having one of

CC 12614 nucleic acid sequences mentioned in the specification, or their

CC complements or the 12387 open reading frames derived from the 12614

CC probes. Also included are a microarray comprising the novel set of

CC probes; the novel set of probes which hybridise at high stringency to a

CC nucleic acid expressed in the human lung; measuring gene expression in a

CC sample derived from human lung, comprising (a) contacting the array with

CC a collection of detectably labeled nucleic acids derived from human lung

CC mRNA, and (b) measuring the label detectably bound to each probe of

CC the array; identifying exons in a eukaryotic genome, comprising

CC (a) algorithmically predicting at least one exon from genomic sequences

CC of the eukaryote; and (b) detecting specific hybridisation of detectably

CC labeled nucleic acids from eukaryote lung mRNA, to a single exon probe,

CC having a fragment identical to the predicted exon, the probe is included

CC in the above mentioned microarray; assigning exons to a single gene,

CC comprising (a) identifying exons from genomic sequence by the method

CC above and (b) measuring the expression of each of the exons in several

CC tissues and/or cell types using hybridisation to a single exon

CC microarrays having a probe with the exon, where a common pattern of

CC expression of the exons in the tissues and/or cell types indicates that

CC the exons should be assigned to a single gene; a peptide comprising one

CC of 12011 sequences, mentioned in the specification, or encoded by the

CC probes/open reading frames (ORF). The probes are used for gene

CC expression analysis, and for identifying exons in a gene, particularly

CC using human lung derived mRNA and for the study of lung diseases

CC such as asthma, lung cancer, chronic obstructive pulmonary disease

CC (COPD), interstitial lung disease (ILD), familial idiopathic pulmonary

CC fibrosis, neurofibromatosis, tuberous sclerosis, Gaucher's disease,

CC Niemann-Pick disease, Hermansky-Pudlak syndrome, sarcoidosis, pulmonary

CC haemostasis, pulmonary histiocytosis, lymphangioleiomyomatosis,

CC pulmonary alveolar proteinosis, Karagener syndrome, fibrocystic

CC pulmonary dysplasia, primary ciliary dyskinesia, pulmonary hypertension

CC and hyaline membrane disease. The present sequence is a peptide/protein

CC encoded by a single exon probe of the invention.

CC Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic

CC format directly from WIPD at

CC ftp.wipo.int/pub/published_pct_sequences.

XX

XX Sequence 153 AA;

SQ

QY 293 GAAAGCAA 301

DB 56 GAAAGCAA 64

Query Match 1.8%; Score 9; DB 23; Length 153;

Best Local Similarity 100.0%; Pred. No. 8.1;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 31

AAAG1693

ID AAAG1693 standard; Protein: 159 AA.

XX AAAG1693;

XX 17-OCT-2000 (first entry)

DE Arabidopsis thaliana protein fragment SEQ ID NO: 10507.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

OS Arabidopsis thaliana.

XX

PN EP1033405-A2.

XX

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-0301439.

XX

PR 25-FEB-1999; 990S-0121825.

XX 05-MAR-1999; 990S-0123180.

PR 09-MAR-1999; 990S-0123548.

PR 23-MAR-1999; 990S-0125788.

PR 25-MAR-1999; 990S-0126264.

PR 29-MAR-1999; 990S-0126785.

PR 01-APR-1999; 990S-0127462.

PR 06-APR-1999; 990S-0128234.

PR 08-APR-1999; 990S-0128714.

PR 16-APR-1999; 990S-0129845.

PR 19-APR-1999; 990S-0130077.

PR 21-APR-1999; 990S-0130449.

PR 23-APR-1999; 990S-0130510.

PR 23-APR-1999; 990S-0130891.

PR 28-APR-1999; 990S-0131449.

PR 30-APR-1999; 990S-0132048.

PR 30-APR-1999; 990S-0132407.

PR 04-MAY-1999; 990S-0132484.

PR 05-MAY-1999; 990S-0132485.

PR 06-MAY-1999; 990S-0132486.

PR 06-MAY-1999; 990S-0132487.

PR 07-MAY-1999; 990S-0132863.

XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488900/53.
DR Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human bone marrow -
PT Example 4; SEQ ID NO: 32152; 658bp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC bone marrow. They can be used to measure gene expression in bone marrow
CC samples, which may enable the improved diagnosis and treatment of cancers
CC such as lymphoma, leukemia and myeloma. The present sequence is a
CC protein encoded by one of the probes of the invention.
XX
SQ Sequence 153 AA;
Query Match 1.8%; Score 9; DB 22; Length 153;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 293 GAAAGGAA 301
DB 56 GAAAGGAA 64
RESULT 28
AAM19335
ID AAM19335 standard; Protein; 153 AA.
XX AAM19335;
XX 12-OCT-2001 (first entry)
DT Peptide #5769 encoded by probe for measuring cervical gene expression.
XX
DE Probe; human; microarray; gene expression; cervical epithelial cell;
KW cervical cancer.
XX
XX Homo sapiens.
OS
PN WO200157278-A2.
PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00670.
PE
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488901/53.
DR Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human cervical epithelial cells -
PT
XX Claim 27; SEQ ID NO 24161; 487bp; English.
XX
XX The present invention relates to human single exon nucleic acid probes
CC (SEN: see AAI10068-AA128459). The present sequence is a peptide encoded
CC by one such probe. The SENs are derived from human HeLa cells. The SENs
CC can be used to produce a single exon microarray, which can be used for
CC measuring human gene expression in a sample derived from human cervical

CC epithelial cells. By measuring gene expression, the probes are therefore
CC useful in grading and/or staging of diseases of the cervix, notably
CC cervical cancer.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 153 AA;
Query Match 1.8%; Score 9; DB 22; Length 153;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 293 GAAAGGAA 301
DB 56 GAAAGGAA 64
RESULT 29
AAM32131
ID AAM32131 standard; Protein; 153 AA.
XX AAM32131;
XX 17-OCT-2001 (first entry)
DT Peptide #6168 encoded by probe for measuring placental gene expression.
XX
DE Probe; microarray; human; placenta; antenatal diagnosis;
KW genetic disorder.
XX
XX Homo sapiens.
OS
PN WO200157272-A2.
PD 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00663.
PE
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-488897/53.
DR Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human placenta -
PT
XX Claim 27; SEQ ID NO 32400; 654bp; English.
XX
XX The present invention relates to single exon nucleic acid probes (SEN:
CC see AAI13135-AA157546). The present sequence is a peptide encoded by one
CC such probe. The probes are useful for producing a microarray for
CC predicting, measuring and displaying gene expression in samples derived
CC from human placenta. The probes are useful for antenatal diagnosis of
CC human genetic disorders.
XX
SQ Sequence 153 AA;
Query Match 1.8%; Score 9; DB 22; Length 153;
Best Local Similarity 100.0%; Pred. No. 8.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 293 GAAAGGAA 301
DB 56 GAAAGGAA 64

XX Human; gene expression; heart; microarray; vascular system;
KM cardiovascular disease; hypertension; cardiac arrhythmia;
KM congenital heart disease.
XX
OS Homo sapiens.
XX
PN WO200157274-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00666.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-488899/53.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT hearts -
PS
XX Claim 15; SEQ ID No 25516; 530pp; English.
XX
CC The present invention relates to single exon nucleic acid probes for
CC measuring human gene expression in a sample derived from human heart (see
CC ABA21533-ABA41305). The present sequence is a protein encoded by one such
CC probe. The probes may be used for predicting, measuring and displaying
CC gene expression in samples derived from the human heart via microarrays.
CC By measuring gene expression, the probes are useful for predicting,
CC diagnosing, grading, staging, monitoring and prognosing diseases of the
CC human heart and vascular system e.g. cardiovascular disease,
CC hypertension, cardiac arrhythmias and congenital heart disease.
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 153 AA;
XX
XX Query Match 1.8%; Score 9; DB 22; Length 153;
XX Best Local Similarity 100.0%; Pred. No. 8.1;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 293 GAAAGCAA 301
XX |||||||||
DB 56 GAAAGCAA 64
XX
XX
XX RESULT 26
XX ID AAM59299 standard; Protein; 153 AA.
XX
XX AAM59299;
XX
XX 05-NOV-2001 (first entry)
XX
XX Human brain expressed single exon probe encoded protein SEQ ID NO: 31404.
XX
XX Human; brain expressed exon; gene expression analysis; probe;
KM microarray; Alzheimer's disease; multiple sclerosis; schizophrenia;
KM epilepsy; cancer.
XX
XX Homo sapiens.
XX
XX WO200157275-A2.
XX
XX

XX 09-AUG-2001.
PD
XX 30-JAN-2001; 2001WO-US00667.
PF
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX Penn SG, Hanzel DK, Chen W, Rank DR;
PI
XX WPI; 2001-483446/52.
XX
XX Single exon nucleic acid probes for analyzing gene expression in human
PT brains -
PS
XX Example 4; SEQ ID NO: 31404; 650pp + Sequence Listing; English.
XX
XX The present invention provides a number of single exon nucleic acid
CC probes which are derived from genomic sequences expressed in the human
CC brain. They can be used to measure gene expression in brain cell samples,
CC which may enable the diagnosis and improved treatment of nervous system
CC diseases such as Alzheimer's disease, multiple sclerosis, schizophrenia,
CC epilepsy and cancers. The present sequence is a protein encoded by one of
CC the probes of the invention.
XX
SQ Sequence 153 AA;
XX
XX Query Match 1.8%; Score 9; DB 22; Length 153;
XX Best Local Similarity 100.0%; Pred. No. 8.1;
XX Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
OY 293 GAAAGCAA 301
XX |||||||||
DB 56 GAAAGCAA 64
XX
XX
XX RESULT 27
XX ID AAM71846 standard; Protein; 153 AA.
XX
XX AAM71846;
XX
XX 06-NOV-2001 (first entry)
XX
XX Human bone marrow expressed probe encoded protein SEQ ID NO: 32152.
XX
XX Human; bone marrow expressed exon; gene expression analysis; probe;
KM microarray; cancer; leukemia; lymphoma; myeloma.
XX
XX Homo sapiens.
XX
XX WO200157276-A2.
XX
XX 09-AUG-2001.
XX
XX 30-JAN-2001; 2001WO-US00668.
XX
XX 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX
XX

```

PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148365.
PR 16-AUG-1999; 99US-0148684.
PR 17-AUG-1999; 99US-0149368.
PR 18-AUG-1999; 99US-0149175.
PR 20-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 23-AUG-1999; 99US-0149829.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0149930.
PR 26-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151080.
PR 31-AUG-1999; 99US-0151303.
PR 01-SEP-1999; 99US-0151438.
PR 07-SEP-1999; 99US-0151930.
PR 10-SEP-1999; 99US-0152363.
PR 13-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 16-SEP-1999; 99US-0154018.
PR 20-SEP-1999; 99US-0154039.
PR 22-SEP-1999; 99US-0154779.
PR 23-SEP-1999; 99US-0155139.
PR 24-SEP-1999; 99US-0155486.
PR 28-SEP-1999; 99US-0156459.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161404.
PR 23-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

```

```

Query Match 1.8%; Score 9; DB 21; Length 149;
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 293 GAAAAGCAA 301

```

```

DB 119 GAAAAGCAA 127
|||||
RESULT 24
ABR38670
ID ABR38670 standard; Peptide; 153 AA.
ABR38670;
AC ABR38670;
DT 04-FEB-2002 (first entry)
DE Peptide #6176 encoded by human foetal liver single exon probe.
XX Human; foetal liver; gene expression; single exon nucleic acid probe.
XX Homo sapiens.
OS WO200157277-A2.
PN 09-AUG-2001.
PD 30-JAN-2001; 2001WO-US00669.
PF 04-FEB-2000; 2000US-0180312.
PR 26-MAY-2000; 2000US-0207456.
PR 30-JUN-2000; 2000US-0608408.
PR 03-AUG-2000; 2000US-0632366.
PR 21-SEP-2000; 2000US-0234687.
PR 27-SEP-2000; 2000US-0236359.
PR 04-OCT-2000; 2000GB-0024263.
XX (MOLE-) MOLECULAR DYNAMICS INC.
XX Penn SG, Hanzel DK, Chen W, Rank DR;
XX WPI; 2001-483447/52.
XX
XX Human genome-derived single exon nucleic acid probes useful for
XX analyzing gene expression in human fetal liver -
XX
XX Claim 27: SEQ ID NO 31305; 639pp + sequence listing; English.
XX
XX The invention relates to a single exon nucleic acid probe for
XX measuring human gene expression in a sample derived from human foetal
XX liver. The single exon nucleic acid probes may be used for predicting,
XX measuring and displaying gene expression in samples derived from human
XX fetal liver. The present sequence is a peptide encoded by a single exon
XX nucleic acid probe of the invention.
XX Note: The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pat_sequences.
XX
XX Sequence 153 AA:
SQ
QY 293 GAAAAGCAA 301
|||||
DB 56 GAAAAGCAA 64
RESULT 25
ABR23746
ID ABR23746 standard; Protein; 153 AA.
XX
XX ABR23746;
AC
XX
XX DT 23-JAN-2002 (first entry)
XX
XX Protein #5745 encoded by probe for measuring heart cell gene expression.

```

PR 29-OCT-1999; 99US-0162142.
Query Match 1.8%; Score 9; DB 21; Length 118;
Best Local Similarity 100.0%; Fred. No. 6.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 293 GAAAGGAA 301
| | | | | | | |
DB 88 GAAAGGAA 96
RESULT 23
AAG11694
ID AAG11694 standard; Protein; 149 AA.
XX AAG11694;
AC AAG11694;
XX 17-OCT-2000 (first entry)
DT 17-OCT-2000 (first entry)
XX Arabidopsis thaliana protein fragment SEQ ID NO: 10508.
DE Arabidopsis thaliana protein fragment SEQ ID NO: 10508.
XX
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX
XX EPI03405-A2.
PN EPI03405-A2.
XX
XX 06-SEP-2000.
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-0301439.
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130409.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139452.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.

PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 21-JUN-1999; 99US-0139763.
PR 22-JUN-1999; 99US-0139817.
PR 23-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140693.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142380.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144332.
PR 20-JUL-1999; 99US-0144632.
PR 21-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 23-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153700.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 18-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 22-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 23-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 25-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.

PR 05-FEB-1999; 99GB-0002555.
XX (NEUT-) NEUTEC PHARMA PLC.
XX
XX Burnie JP, Matthews RC;
XX
XX WPI: 2000-543485/49.
XX
XX
XX New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,
PT preventing and treating C. pneumoniae infection and atherosclerosis,
PT including coronary atherosclerosis -
PS Claim 5; Page 34; 35pp; English.
XX
XX AAB08324-36 represent epitopes derived from a Chlamydia pneumoniae
CC protein. The protein, immunogenic fragments of it, nucleotide sequences
CC encoding it, or inhibitor specific against it are used to manufacturing
CC a medicament for the treatment of infection due to C. pneumoniae.
CC An antibody specific against the protein can diagnose a C. pneumoniae
CC infection. C. pneumoniae infection can be prevented. Atherosclerosis,
CC including coronary atherosclerosis, caused by C. pneumoniae can also
CC be prevented or treated.
XX
XX
SQ Sequence 9 AA;
Query Match 1.8%; Score 9; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 235 EYDEMKAEE 243
Db 1 EYDEMKAEE 9
RESULT 21
AAR72669
ID AAR72669 standard; Protein; 111 AA.
XX
XX AAR72669;
AC
DT 25-OCT-1995 (first entry)
XX
XX Cladosporium herbarum allergen Clah1.
DE
XX Cladosporium herbarum allergen Clah1.
KM Fungal spore; allergen; Clah1; allergy; ribosomal protein; RL42.
OS
XX Cladosporium herbarum.
PN
XX W09506121-A.
XX
XX 02-MAR-1995.
PD
XX 24-AUG-1994; 94MO-AT00120.
PF
XX 27-AUG-1993; 93AT-0001725.
PR
XX (BIOM-) BIOMAY PRODN & HANDELSGES MBH.
PA
XX Achatz G, Breitenbach M, Ebner C, Hirschwehr R;
PI Kraft D, Lechenauer E, Oberkofler H, Prillinger H;
PI Simon B, Unger A;
XX
XX WPI: 1995-106850/14.
DR N-PSDB; AA087846.
XX
XX Allergens derived from Cladosporium herbarum spores - also
PT recombinant DNA for expressing the allergens, useful for in vitro
PT allergy detection
XX
XX Claim 1; Page 25; 35pp; German.
PS
XX Spores of Cladosporium herbarum are the most common fungal spores
CC found in the air; they can cause allergic reactions. Various Clah
CC

CC allergens and sequences encoding them have now been isolated. The
CC mature Clah1 allergen has mol. wt. 11 kD and is encoded by cDNA
CC sequence AA087846. The allergen has homology to the ribosomal protein
CC RL42. Potential epitopic subfragments were identified by computer
CC analysis of the amino acid sequence. See AAR72670-R72674 for potential
CC B-cell epitopes and AAR72675-R72677 for potential T-cell epitopes.
XX
XX
SQ Sequence 111 AA;
Query Match 1.8%; Score 9; DB 16; Length 111;
Best Local Similarity 100.0%; Pred. No. 6.1;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 293 GAAAGGAA 301
Db 77 GAAAGGAA 85
RESULT 22
AAG11695
ID AAG11695 standard; Protein; 118 AA.
XX
XX AAG11695;
AC
DT 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana protein fragment SEQ ID NO: 10509.
DE
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
XX Arabidopsis thaliana.
OS
XX
XX EP1033405-A2.
PN
XX
XX 06-SEP-2000.
PD
XX
XX 25-FEB-2000; 2000EP-0301439.
PF
XX 25-FEB-1999; 9905-0121825.
PR 05-MAR-1999; 9905-0123180.
PR 09-MAR-1999; 9905-0123548.
PR 23-MAR-1999; 9905-0125788.
PR 25-MAR-1999; 9905-0126264.
PR 29-MAR-1999; 9905-0126785.
PR 01-APR-1999; 9905-0127462.
PR 06-APR-1999; 9905-0128234.
PR 08-APR-1999; 9905-0128714.
PR 16-APR-1999; 9905-0129845.
PR 19-APR-1999; 9905-0130077.
PR 21-APR-1999; 9905-0130449.
PR 23-APR-1999; 9905-0130510.
PR 28-APR-1999; 9905-0130891.
PR 30-APR-1999; 9905-0131449.
PR 30-APR-1999; 9905-0132048.
PR 04-MAY-1999; 9905-0132407.
PR 05-MAY-1999; 9905-0132484.
PR 06-MAY-1999; 9905-0132485.
PR 06-MAY-1999; 9905-0132486.
PR 07-MAY-1999; 9905-0132487.
PR 11-MAY-1999; 9905-0132863.
PR 14-MAY-1999; 9905-0134256.
PR 14-MAY-1999; 9905-0134218.
PR 14-MAY-1999; 9905-0134219.
PR 14-MAY-1999; 9905-0134221.
PR 14-MAY-1999; 9905-0134370.
PR 18-MAY-1999; 9905-0134768.
PR 19-MAY-1999; 9905-0134941.
PR 20-MAY-1999; 9905-0135124.
PR 21-MAY-1999; 9905-0135353.
PR 24-MAY-1999; 9905-0135629.
PR 25-MAY-1999; 9905-0136021.

CC infection. C. pneumoniae infection can be prevented. Artherosclerosis,
 CC including coronary atherosclerosis, caused by C. pneumoniae can also
 CC be prevented or treated.
 XX

SO Sequence 9 AA;

Query Match 1.8%; Score 9; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 175 GVTPESEVI 183
 DB 1 GVTPESEVI 9

RESULT 18
 AAB08330

ID AAB08330 standard; Peptide: 9 AA.

XX AAB08330;

XX 04-DEC-2000 (first entry)

DE Epitope derived from a hlamydia pneumoniae protein.

XX Chlamydia pneumoniae protein; infection; artherosclerosis;

KW coronary atherosclerosis; epitope.

OS Chlamydia pneumoniae.

XX WO200046359-A2.

XX 10-AUG-2000.

PF 28-JAN-2000; 2000WO-GB00237.

XX 05-FEB-1999; 99GB-0002555.

XX (NEUT-) NEUTEC PHARMA PLC.

XX Burnie JP, Mathews RC;

XX WPI: 2000-543485/49.

XX New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,
 PT preventing and treating C. pneumoniae infection and atherosclerosis,
 PT including coronary atherosclerosis -

XX Claim 5; Page 34; 35pp; English.

CC AAB08324-36 represent epitopes derived from a Chlamydia pneumoniae
 CC protein. The protein, immunogenic fragments of it, nucleotide sequences
 CC encoding it, or inhibitor specific against it are used to manufacturing
 CC a medicament for the treatment of infection due to C. pneumoniae.
 CC An antibody specific against the protein can diagnose a C. pneumoniae
 CC infection. C. pneumoniae infection can be prevented. Artherosclerosis,
 CC including coronary atherosclerosis, caused by C. pneumoniae can also
 CC be prevented or treated.
 XX

SO Sequence 9 AA;

Query Match 1.8%; Score 9; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 37 SGNETKQIQ 45
 DB 1 SGNETKQIQ 9

RESULT 19

XX AAB08333 standard; Peptide: 9 AA.

XX

AAB08333;

XX 04-DEC-2000 (first entry)

DE Epitope derived from a hlamydia pneumoniae protein.

XX Chlamydia pneumoniae protein; infection; artherosclerosis; †

KW coronary atherosclerosis; epitope.

OS Chlamydia pneumoniae.

XX WO200046359-A2.

XX 10-AUG-2000.

PF 28-JAN-2000; 2000WO-GB00237.

XX 05-FEB-1999; 99GB-0002555.

XX (NEUT-) NEUTEC PHARMA PLC.

XX Burnie JP, Mathews RC;

XX WPI: 2000-543485/49.

XX New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,
 PT preventing and treating C. pneumoniae infection and atherosclerosis,
 PT including coronary atherosclerosis -

XX Claim 5; Page 34; 35pp; English.

CC AAB08324-36 represent epitopes derived from a Chlamydia pneumoniae
 CC protein. The protein, immunogenic fragments of it, nucleotide sequences
 CC encoding it, or inhibitor specific against it are used to manufacturing
 CC a medicament for the treatment of infection due to C. pneumoniae.
 CC An antibody specific against the protein can diagnose a C. pneumoniae
 CC infection. C. pneumoniae infection can be prevented. Artherosclerosis,
 CC including coronary atherosclerosis, caused by C. pneumoniae can also
 CC be prevented or treated.
 XX

SO Sequence 9 AA;

Query Match 1.8%; Score 9; DB 21; Length 9;
 Best Local Similarity 100.0%; Pred. No. 7.8e+05;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 127 SKSMESTLE 135
 DB 1 SKSMESTLE 9

RESULT 20

XX AAB08334 standard; Peptide: 9 AA.

XX AAB08334;

XX 04-DEC-2000 (first entry)

DE Epitope derived from a hlamydia pneumoniae protein.

XX Chlamydia pneumoniae protein; infection; artherosclerosis;

KW coronary atherosclerosis; epitope.

OS Chlamydia pneumoniae.

XX WO200046359-A2.

XX 10-AUG-2000.

PF 28-JAN-2000; 2000WO-GB00237.


```

XX 21-APR-2000; 2000US-199047P.
PR 02-JUN-2000; 2000US-208841P.
PR 07-JUL-2000; 2000US-216747P.
XX
PA (CORI-) CORIXA CORP.
XX
PI Skelky YAM, Persing DH, Mitcham JL, Wang SS, Bhatia A;
PI L'maisonneuve J, Zhang Y, Jen S, Carter D;
XX
DR WPI; 2001-616774/71.
DR N-PSDB; AAS59546.
XX
PT Propionibacterium acnes polypeptides and nucleic acids useful for
PT vaccinating against and diagnosing infections, especially useful for
PT treating acne vulgaris -
XX
XX Example 1; SEQ ID No 11250; 1069pp; English.
XX
CC Sequences AAU39105-AAU68017 represent Propionibacterium acnes immunogenic
CC polypeptides. The proteins and their associated DNA sequences are used in
CC the treatment, prevention and diagnosis of medical conditions caused by
CC P. acnes. The disorders include SAPHO syndrome (synovitis, acne,
CC pustulosis, hyperostosis and osteomyelitis), uveitis and endophthalmitis.
CC P. acnes is also involved in infections of bone, joints and the central
CC nervous system, however it is particularly involved in the inflammatory
CC lesions associated with acne vulgaris. A method for detecting the
CC presence or absence of P. acnes in a patient comprises contacting a
CC sample with a binding agent that binds to the proteins of the invention
CC and determining the amount of bound protein in the sample. The
CC polypeptides may be used as antigens in the production of antibodies
CC specific for P. acnes proteins. These antibodies can be used to
CC downregulate expression and activity of P. acnes polypeptides and
CC therefore treat P. acnes infections. The antibodies may also be used as
CC diagnostic agents for determining P. acnes presence, for example, by
CC enzyme linked immunosorbent assay (ELISA).
CC Note: The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 755 AA:

```

Query Match 2.0%; Score 10; DB 22; Length 755;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 281 AGLAGLAAGA 290
   |||||
DB 575 AGLAGLAAGA 584

```

RESULT 16
AAB08324
ID AAB08324 standard; Peptide; 9 AA.
XX
AC AAB08324;
XX
DT 04-DEC-2000 (first entry)
XX
DE Epitope derived from a hlamydia pneumoniae protein.
XX
KM Chlamydia pneumoniae protein: infection; atherosclerosis;
KM coronary atherosclerosis; epitope.
XX
OS Chlamydia pneumoniae.
XX
PN WO200046359-A2.
XX
PD 10-AUG-2000.
XX
PF 28-JAN-2000; 2000MO-GB00237.
XX
PR 05-FEB-1999; 99GB-0002555.
XX

```

XX (NEUT-) NEUTEC PHARMA PLC.
XX
PA Burnie JP, Matthews RC;
XX
PI WPI; 2000-543485/49.
XX
DR WPI; 2000-543485/49.
XX
XX
XX New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,
XX preventing and treating C. pneumoniae infection and atherosclerosis,
XX including coronary atherosclerosis -
XX
XX Claim 5; Page 32; 35pp; English.
XX
CC AAB08324-36 represent epitopes derived from a Chlamydia pneumoniae
CC protein. The protein, immunogenic fragments of it, nucleotide sequences
CC encoding it, or inhibitor specific against it are used to manufacturing
CC a medicament for the treatment of infection due to C. pneumoniae.
CC An antibody specific against the protein can diagnose a C. pneumoniae
CC infection. C. pneumoniae infection can be prevented. Atherosclerosis,
CC including coronary atherosclerosis, caused by C. pneumoniae can also
CC be prevented or treated.
XX
XX
SQ Sequence 9 AA:

```

Query Match 1.8%; Score 9; DB 21; Length 9;
Best Local Similarity 100.0%; Pred. No. 7.8e+05;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 163 SAKLEPPEL 171
   |||||
DB 1 SAKLEPPEL 9

```

RESULT 17
AAB08326
ID AAB08326 standard; Peptide; 9 AA.
XX
AC AAB08326;
XX
DT 04-DEC-2000 (first entry)
XX
DE Epitope derived from a hlamydia pneumoniae protein.
XX
KM Chlamydia pneumoniae protein: infection; atherosclerosis;
KM coronary atherosclerosis; epitope.
XX
OS Chlamydia pneumoniae.
XX
PN WO200046359-A2.
XX
PD 10-AUG-2000.
XX
PF 28-JAN-2000; 2000MO-GB00237.
XX
PR 05-FEB-1999; 99GB-0002555.
XX
PA (NEUT-) NEUTEC PHARMA PLC.
XX
PI Burnie JP, Matthews RC;
XX
DR WPI; 2000-543485/49.
XX
PT New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,
PT preventing and treating C. pneumoniae infection and atherosclerosis,
PT including coronary atherosclerosis -
XX
XX Claim 5; Page 33; 35pp; English.
XX
CC AAB08324-36 represent epitopes derived from a Chlamydia pneumoniae
CC protein. The protein, immunogenic fragments of it, nucleotide sequences
CC encoding it, or inhibitor specific against it are used to manufacturing
CC a medicament for the treatment of infection due to C. pneumoniae.
CC An antibody specific against the protein can diagnose a C. pneumoniae
CC infection. C. pneumoniae infection can be prevented. Atherosclerosis,
CC including coronary atherosclerosis, caused by C. pneumoniae can also
CC be prevented or treated.
XX

XX	Leach MD, Shinkets RA;
PI	
XX	WPI: 2002-106200/14.
DR	N-PSDB; ABBN75498.
XX	
PT	Novel human polypeptides and polynucleotides useful for diagnosing,
PT	preventing and treating cardiovascular disease, neurodegenerative,
PT	hyperproliferative disorders and disorders related to organ
PT	transplantation
XX	
PS	Claim 10; Page 491; 2508pp; English.
XX	
CC	Sequences ABP31028-ABP35561 represent 4534 novel human proteins
CC	designated ORF (open reading frame) 1-4534, and sequences ABBN75054-
CC	ABBN79587 represent cDNAs encoding them. The invention also encompasses
CC	polypeptides at least 808 identical to the ORF1-ORF4534 (collectively
CC	referred to as ORX) proteins, polynucleotides at least 85% identical to
CC	the ORX nucleic acid sequences, vectors and host cells comprising ORX
CC	polynucleotides, the recombinant production of ORX proteins, antibodies
CC	specific for ORX proteins, methods of detecting ORX polynucleotides and
CC	polypeptides, methods of screening for modulators of ORX expression or
CC	activity, and methods of screening individuals for a predisposition to an
CC	ORX-associated disorder. The ORX proteins of the invention have a wide
CC	range of biological activities, such as cytokine, cell proliferation,
CC	cell differentiation, immune modulation, haematopoiesis regulation,
CC	tissue growth, angiogenesis, activin or inhibin activity, chemotactic/
CC	chemokinetic activity, haemostatic activity, thrombolytic activity,
CC	receptor/ligand, antiinflammatory activity, tumour inhibition activity,
CC	and antinfecitive activity, and may also be involved in the determination
CC	of bodily characteristics, fertility and behaviour. ORX proteins,
CC	nucleic acids and antibodies may be used in the treatment of cancers,
CC	other proliferative disorders such as psoriasis and benign tumours,
CC	neurological disorders such as epilepsy and Alzheimer's disease,
CC	cardiovascular diseases, immune system disorders, disorders related to
CC	organ transplantation, disorders of tissue growth and regeneration, to
CC	diseases such as diabetes mellitus, hypothyroidism, and cholesterol
CC	storage disease, and infectious diseases caused by viral, bacterial,
CC	fungal and other pathogens. ORX nucleic acids may also be used as a
CC	source of primers and probes, in the detection of ORX genomic sequences
CC	or transcripts, in the identification and cloning of homologous
CC	sequences, in genetic diagnosis, and in forensic biology. The ORX
CC	nucleic acids may additionally be used to produce transgenic animals
CC	which may be useful for studying the function and/or activity of ORX
CC	protein, and in drug screening. The ORX proteins may also be used as
CC	immunogens to generate specific antibodies, which are useful in the
CC	diagnosis, treatment and monitoring of ORX-associated diseases.
XX	
SO	Sequence 129 AA;
SO	
QY	281 AGIAGIAGA 290
Db	77 AGLAGIAGA 86
XX	
RESULT 14	
ID	AAO14994
XX	AAO14994 standard; Protein: 720 AA.
AC	AAO14994;
DT	06-AUG-2002 (first entry)
XX	
DE	Laminin-related protein.
XX	
KW	laminin gamma-2; cancer; laminin gamma-2 chain inhibition;
KW	carcinogen inhibition; anti-gamma-2 chain antibody;
KW	epithelial cell adhesion; laminin-5.
XX	

OS	Unidentified.
XX	
PN	US2002052307-A1.
XX	
PD	02-MAY-2002.
XX	
PE	08-JAN-2001; 2001US-0756071.
XX	
PR	07-JAN-2000; 2000US-175005P.
XX	
PR	04-OCT-1994; 94DS-0317450.
XX	
PR	18-FEB-1997; 97US-0800593.
XX	
PR	15-SEP-2000; 2000US-0663147.
XX	
PA	(TRYG/) TRYGVASON K.
XX	
PA	(KALL/) KALLUNKI P.
XX	
PI	(PYKE/) PYKE C.
XX	
Pt	Trygvason K, Kallunki P, Pyke C;
DR	WPI; 2002-434824/46.
XX	
PT	Modulating laminin 5 gamma 2 chain interactions of invasive carcinogens
XX	
PT	for treating cancers and promoting attachment of cultured cells in
XX	
PT	vitro -
XX	
PS	Disclosure; Page 34-36; 51pp; English.
XX	
CC	The invention comprises a method of inhibiting the laminin gamma-2 chain
XX	
CC	interactions of invasive carcinogens with surrounding tissues - by using
XX	
CC	anti-gamma-2 chain antibodies to inhibit the gamma-2 chain biological
XX	
CC	activity of the invasive carcinogens. The invention also comprises a
XX	
CC	method for promoting adhesion of epithelial cells by exposing the cells
XX	
CC	to intact laminin-5 molecules. The first method of the invention is
XX	
CC	useful for preventing gamma 2 chain interactions of invasive carcinogens
XX	
CC	with surrounding tissues. The second method of the invention is useful
XX	
CC	for promoting adhesion of cultured epithelial cells.
XX	
CC	NOTE: The present sequence is shown as a protein in the sequence listing,
XX	
CC	however in a figure of the invention this sequence is shown to be a DNA
XX	
CC	sequence (see AAL42916).
XX	
SO	Sequence 720 AA:
QY	293 GAAAGGAG 302
DB	602 GAAAGGAG 611
Query Match	2.0%; Score 10; DB 23; Length 720;
Best Local Similarity	100.0%; Pred. No. 4.1;
Matches 10; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
RESULT 15	
AAU50055	
ID	AAU50055 standard; Protein; 755 AA.
AC	AAU50055;
XX	
DT	27-FEB-2002 (first entry)
XX	
DE	Propionibacterium acnes immunogenic protein #10951.
SAPHO syndrome; synovitis; acne; pustulosis; hyperostosis; osteomyelitis;	
uveitis; endophthalmitis; bone; joint; central nervous system; ELISA;	
Inflammatory lesion; acne vulgaris; enzyme linked immunosorbent assay;	
dermatological; osteopathic; neuroprotectant.	
Propionibacterium acnes.	
MO200181581-A2.	
01-NOV-2001.	
20-APR-2001; 2001WO-US12865.	

AC AAB08336;
 XX
 DF 04-DEC-2000 (first entry)
 XX
 DE Epitope derived from a hlamydia pneumoniae protein.
 XX
 KW Chlamydia pneumoniae protein; infection; atherosclerosis;
 KW coronary atherosclerosis; epitope.
 XX
 OS Chlamydia pneumoniae.
 XX
 PN WO200046359-A2.
 XX
 PD 10-AUG-2000.
 XX
 PF 28-JAN-2000; 2000WO-GB00237.
 XX
 PR 05-FEB-1999; 99GB-0002555.
 XX
 PA (NEUT-) NEUTEC PHARMA PLC.
 XX
 PI Burnie JP, Matthews RC;
 XX
 DR WPI; 2000-543485/49.
 XX
 PR New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,
 PR preventing and treating C. pneumoniae infection and atherosclerosis,
 PR including coronary atherosclerosis -
 XX
 PS Claim 5; Page 35; 35pp; English.
 XX
 CC AAB08334-36 represent epitopes derived from a Chlamydia pneumoniae
 CC protein. The protein, immunogenic fragments of it, nucleotide sequences
 CC encoding it, or inhibitor specific against it are used to manufacturing
 CC a medicament for the treatment of infection due to C. pneumoniae.
 CC An antibody specific against the protein can diagnose a C. pneumoniae
 CC infection. C. pneumoniae infection can be prevented. Atherosclerosis,
 CC including coronary atherosclerosis, caused by C. pneumoniae can also
 CC be prevented or treated.
 XX
 SQ Sequence 14 AA;
 XX
 Query Match 2.8%; Score 14; DB 21; Length 14;
 Best Local Similarity 100.0%; Pred. No. 3.5e-05;
 Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 167 ETPELPKPGVTPRS 180
 DB 1 ETPELPKPGVTPRS 14
 XX
 RESULT 12
 ABB58985
 ID ABB58985 standard; Protein; 1937 AA.
 XX
 AC ABB58985;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 3747.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US09231.
 XX
 PR 23-MAR-2000; 2000US-191637P.
 XX

PR 11-JUL-2000; 2000US-0614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL03088.
 XX
 PR New isolated nucleic acid detection reagent for detecting 1000 or more
 PR genes from Drosophila and for elucidating cell signalling and cell-cell
 PR interactions -
 XX
 PS Disclosure; SEQ ID NO 3747; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins
 CC (ABB57737-ABB72072).
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic format directly from WIPO
 CC at ftp.wipo.int/pub/published_pcl_sequences.
 XX
 SQ Sequence 1937 AA;
 XX
 Query Match 2.2%; Score 11; DB 22; Length 1937;
 Best Local Similarity 100.0%; Pred. No. 1.2;
 Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 294 AAAAGCAAGCA 304
 DB 343 AAAAGCAAGCA 353
 XX
 RESULT 13
 ABB31472
 ID ABB31472 standard; Protein; 129 AA.
 XX
 AC ABB31472;
 XX
 DT 09-JUL-2002 (first entry)
 XX
 DE Human ORF445 protein, SEQ ID NO:890.
 XX
 KW Human; ORF; open reading frame; ORF; drug screening; diagnosis;
 KW disease monitoring; cytokine; cell proliferation; cell differentiation;
 KW immune modulation; haematopoiesis regulation; tissue growth;
 KW angiogenesis; activin; inhibin; chemotactic; chemokine; haemostatic;
 KW thrombolytic; tumour inhibition; bodily characteristic; fertility;
 KW behaviour; cancer; proliferative disorder; neurological disorder;
 KW cardiovascular disease; immune system disorder; organ transplantation;
 KW tissue growth disorder; tissue regeneration disorder; diabetes mellitus;
 KW hypothyroidism; cholesterol ester storage disease; infection; vulnerability;
 KW vasotropic; antipsoriatic; antidiabetic; cytosolic; nocitropic;
 KW neuroprotective; antiatherosclerotic; anticoagulant; thrombolytic;
 KW cardiant; hypotensive; antihypertoid; antiinflammatory; immunomodulator;
 KW dermatological; analgesic; virucide; antibacterial; fungicide.
 XX
 OS Homo sapiens.
 XX
 PN WO200190366-A2.
 XX
 PD 29-NOV-2001.
 XX
 PF 24-MAY-2001; 2001WO-US17076.
 XX
 PR 24-MAY-2000; 2000US-206690P.
 XX
 PA (CURA-) CURAGEN CORP.
 XX

```

AC AAR94585;
XX
XX 11-NOV-1996 (first entry)
XX
DE DHFR/C. pneumoniae antigen variant fusion protein.
XX
XX Polypeptide antigen: strain YK41; plasmid; probe;
XX PCPN533T; primer; assay; detection; antibody; diagnosis;
XX infection; fusion protein; dihydrofolate reductase; DHFR;
XX variant; Chlamydia pneumoniae.
XX
OS Synthetic.
XX
FH Key
FH Protein 1..160
FH FT /label= dihydrofolate reductase
FH FT 162..432
FH FT /label= C. pneumoniae antigen
XX
XX W09609320-A1.
XX
XX 28-MAR-1996.
XX
XX 20-SEP-1995; 95MO-JP01896.
XX
XX 28-APR-1995; 95JP-0106011.
XX 20-SEP-1994; 94JP-0224711.
XX 28-APR-1995; 95JP-0106006.
XX 28-APR-1995; 95JP-0106008.
XX 28-APR-1995; 95JP-0106009.
XX 28-APR-1995; 95JP-0106010.
XX
XX (HITB ) HITACHI CHEM CO LTD.
XX
XX Izutsu H, Matsumoto A, Obara K;
XX WPI; 1996-188399/19.
XX DR N-PSDB; AAT14619.
XX
XX Recombinant Chlamydia pneumoniae antigen and antibodies to it -
XX used for detection and assay of C. pneumoniae e.g. in clinical
XX diagnosis
XX
XX Claim 23; Pages 103-107; 128pp; Japanese.
XX
XX The present sequence is the dihydrofolate reductase (DHFR)/
XX C. pneumoniae polypeptide antigen variant fusion protein.
XX C. pneumoniae strain YK41 was cultured and genomic DNA extracted
XX to prep. a lambda gt11 DNA library. The library was then screened
XX with an anti-YK41 monoclonal antibody (Mab), which was prepd. by
XX fusing spleen cells from a mouse infected with YK41 with myeloma
XX p3/NS1/-A44-1 to produce a Mab expressing hybridoma. The DNA
XX obtd. was then fused with DHFR DNA and the expression vector
XX PAD431 to give PCPN533T. The plasmid was used to transform an
XX E. coli host, which was cultured to give an antigenic polypeptide
XX fusion protein. The fusion protein and primers and probes derived
XX from its DNA can be used in assays for the detection of the
XX antigenic polypeptide antibodies and DNA, respectively, useful in
XX the diagnosis of C. pneumoniae infection.
XX
SQ Sequence 432 AA:
Query Match 52.2%; Score 259; DB 17; Length 432;
Best Local Similarity 100.0%; Pred. No. 3.7e-222;
Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 4 MSISSSGPDNOKNIMSVLTSTPGVPOQDKLSGNETKQIOOTRGKNTMESDATTAG 63
DB 162 MSISSSGPDNOKNIMSVLTSTPGVPOQDKLSGNETKQIOOTRGKNTMESDATTAG 221
QY 64 ASGKDKTSTTETETAPQOGVAAGKSSSESOKAGADTGVSGAAATTTASTATKIAMQTSI 123
DB 222 ASGKDKTSTTETETAPQOGVAAGKSSSESOKAGADTGVSGAAATTTASTATKIAMQTSI 281

```

```

QY 124 EASKSMESTLESTIOSISAQOMEVEAVVVAALSGKSSGSAKLETPELPKPGVTPRSEVI 183
DB 282 EASKSMESTLESTIOSISAQOMEVEAVVVAALSGKSSGSAKLETPELPKPGVTPRSEVI 341
QY 184 EIGIALAKAIQTIQGEATKRSALSNVASTQAOADOTNKLGLEKQAIKIDKEEREYQEKKA 243
DB 342 EIGIALAKAIQTIQGEATKRSALSNVASTQAOADOTNKLGLEKQAIKIDKEEREYQEKKA 401
QY 244 QKSKDLEGIMDVNTVWMA 262
DB 402 QKSKDLEGIMDVNTVWMA 420

RESULT 10
AAB08335
ID AAB08335 standard; Peptide; 14 AA.
XX
XX AAB08335;
XX
XX 04-DEC-2000 (first entry)
XX
XX Epitope derived from a hlamydia pneumoniae protein.
XX
XX Chlamydia pneumoniae protein: infection; arteriosclerosis;
XX coronary atherosclerosis; epitope.
XX
XX Chlamydia pneumoniae.
XX
XX W0200046359-A2.
XX
XX 10-AUG-2000.
XX
XX 28-JAN-2000; 2000WO-GB00237.
XX
XX 05-FEB-1999; 99GB-0002555.
XX
XX (NEUT-) NEUTEC PHARMA PLC.
XX
XX Burnle JP, Matthews RC;
XX
XX WPI; 2000-543485/49.
XX
XX New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,
XX preventing and treating C. pneumoniae infection and arteriosclerosis,
XX including coronary atherosclerosis -
XX
XX Claim 5; Page 34; 35pp; English.
XX
XX AAB08324-36 represent epitopes derived from a Chlamydia pneumoniae
XX protein. The protein, immunogenic fragments of it, nucleotide sequences
XX encoding it, or inhibitor specific against it are used to manufacturing
XX a medicament for the treatment of infection due to C. pneumoniae.
XX An antibody specific against the protein can diagnose a C. pneumoniae
XX infection. C. pneumoniae infection can be prevented. Arteriosclerosis,
XX including coronary atherosclerosis, caused by C. pneumoniae can also
XX be prevented or treated.
XX
SQ Sequence 14 AA:
Query Match 2.8%; Score 14; DB 21; Length 14;
Best Local Similarity 100.0%; Pred. No. 3.5e-05;
Matches 14; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 223 EKQAIKIDKEEREY 236
DB 1 EKQAIKIDKEEREY 14
RESULT 11
AAB08336
ID AAB08336 standard; Peptide; 14 AA.
XX

```

```

XX 20-SEP-1995; 95WO-JP01896.
PF 28-APR-1995; 95JP-0106011.
XX 20-SEP-1994; 94JP-0224711.
PR 20-SEP-1994; 95JP-0106006.
XX 28-APR-1995; 95JP-0106008.
PR 28-APR-1995; 95JP-0106009.
XX 28-APR-1995; 95JP-0106010.
PA (HITB ) HITACHI CHEM CO LTD.
XX Izutsu H, Matsumoto A, Obara K;
PI WPI: 1996-188399/19.
XX N-PSDB; AAT14613.
DR Recombinant Chlamydia pneumoniae antigen and antibodies to it
XX used for detection and assay of C. pneumoniae e.g. in clinical
PT diagnosis
XX Claim 2; Pages 64-66; 128pp; Japanese.
XX The present sequence is a variant of the C. pneumoniae
CC polypeptide antigen, polypeptide A. C. pneumoniae strain YK41 was
CC cultured and genomic DNA extracted to prep. a lambda gt11 DNA
CC library. The library was then screened with an anti-YK41
CC monoclonal antibody (Mab), which was prepd. by fusing spleen cells
CC from a mouse infected with YK41 with myeloma P3/NS1/1-A94-1 to
CC produce a Mab expressing hybridoma. The DNA obtd. was then fused
CC with the expression vector pAD431 to give pCPMS3alpha. The
CC plasmid was used to transform an E. coli host, which was cultured
CC to give the antigenic polypeptide, polypeptide A. Polypeptide A
CC and primers and probes derived from its DNA can be used in assays
CC for the detection of polypeptide A antibodies and DNA,
CC respectively, useful in the diagnosis of C. pneumoniae infection.
CC
XX Sequence 271 AA:
SQ
Query Match 52.2%; Score 259; DB 17; Length 271;
Best Local Similarity 100.0%; Pred. No. 2.3e-222;
Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 4 MSISSSGPDNKNIMSOVLSTPGVPOODKLSGNETKQIOOTROGKNTMESDATTAG 63
DB 1 MSISSSGPDNKNIMSOVLSTPGVPOODKLSGNETKQIOOTROGKNTMESDATTAG 60
OY 64 ASGDKTSTTKTETAPQOGVAAGKESSESOKAGADTGVSGAAATTASNTATKIAMQTSI 123
DB 61 ASGDKTSTTKTETAPQOGVAAGKESSESOKAGADTGVSGAAATTASNTATKIAMQTSI 120
OY 124 EASASMSMESTLESLOSLSAOKMEYEAIVVAALSGKSSGSASAKLELPPELPKPVTRSEYI 183
DB 121 EASASMSMESTLESLOSLSAOKMEYEAIVVAALSGKSSGSASAKLELPPELPKPVTRSEYI 180
OY 184 EIGLALAKAIOTLGEATKSALSNTASTOAOADOTNKLGLEKQAIKIDKREERYOEMKAAE 243
DB 181 EIGLALAKAIOTLGEATKSALSNTASTOAOADOTNKLGLEKQAIKIDKREERYOEMKAAE 240
OY 244 OKSKDLEGTMDTVNTVMIA 262
DB 241 OKSKDLEGTMDTVNTVMIA 259

```

```

XX Chlamydia pneumoniae protein; infection; atherosclerosis;
KW coronary atherosclerosis.
XX Synthetic.
OS Chlamydia pneumoniae.
XX Key Location/Qualifiers
XX Cleavage-site 1..30
XX FT /note= "S-tag and thrombin cleavage site"
XX FT Misc-difference 292..302
XX FT /note= "Histidine tag"
XX WO200046359-A2.
XX 10-AUG-2000.
XX 28-JAN-2000; 2000WO-GB00237.
XX 05-FEB-1999; 99GB-0002555.
XX (NEUT-) NEUTEC PHARMA PLC.
XX Burnie JP, Mathews RC;
XX WPI: 2000-543485/49.
XX New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,
XX preventing and treating C. pneumoniae infection and atherosclerosis,
XX including coronary atherosclerosis -
XX Example; Page 31-32; 35pp; English.
XX The present sequence represents a His-tagged Chlamydia pneumoniae
XX protein. The protein, immunogenic fragments of it, nucleotide sequences
XX encoding it, or inhibitor specific against it are used to manufacturing
XX a medicament for the treatment of infection due to C. pneumoniae.
XX An antibody specific against the protein can diagnose a C. pneumoniae
XX infection. C. pneumoniae infection can be prevented. Atherosclerosis,
XX including coronary atherosclerosis, caused by C. pneumoniae can also
XX be prevented or treated.
XX
XX Sequence 302 AA:
SQ
Query Match 52.2%; Score 259; DB 21; Length 302;
Best Local Similarity 100.0%; Pred. No. 2.7e-222;
Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 1 DTNMSISSSGPDNKNIMSOVLSTPGVPOODKLSGNETKQIOOTROGKNTMESDATTAG 60
DB 33 DTNMSISSSGPDNKNIMSOVLSTPGVPOODKLSGNETKQIOOTROGKNTMESDATTAG 92
OY 61 IAGASGDKTSTTKTETAPQOGVAAGKESSESOKAGADTGVSGAAATTASNTATKIAMQ 120
DB 93 IAGASGDKTSTTKTETAPQOGVAAGKESSESOKAGADTGVSGAAATTASNTATKIAMQ 152
OY 121 TSIEEASMSMESTLESLOSLSAOKMEYEAIVVAALSGKSSGSASAKLELPPELPKPVTRPS 180
DB 153 TSIEEASMSMESTLESLOSLSAOKMEYEAIVVAALSGKSSGSASAKLELPPELPKPVTRPS 212
OY 181 EYIEIGLALAKAIOTLGEATKSALSNTASTOAOADOTNKLGLEKQAIKIDKREERYOEMK 240
DB 213 EYIEIGLALAKAIOTLGEATKSALSNTASTOAOADOTNKLGLEKQAIKIDKREERYOEMK 272
OY 241 AAEQSKDLEGTMDTVNTV 259
DB 273 AAEQSKDLEGTMDTVNTV 291

```

```

RESULT 9
AAB08323
ID AAB08323 standard; Protein: 432 AA.
AC AAB08323;
XX 04-DEC-2000 (first entry)
XX Amino acid sequence of a Chlamydia pneumoniae His-tagged protein.
DE

```

PF 20-SEP-1995; 95WO-JP01896.
 XX 28-APR-1995; 95JP-0106011.
 XX 20-SEP-1994; 94JP-0224711.
 PR 28-APR-1995; 95JP-0106006.
 PR 28-APR-1995; 95JP-0106008.
 PR 28-APR-1995; 95JP-0106009.
 PR 28-APR-1995; 95JP-0106010.
 XX (HITB) HITACHI CHEM CO LTD.
 XX Izutsu H, Matsumoto A, Obara K;
 DR WPI; 1996-188399/19.
 DR N-PSDB; AAT14622.
 XX Recombinant Chlamydia pneumoniae antigen and antibodies to it
 PT used for detection and assay of C. pneumoniae e.g. in clinical
 PT diagnosis
 XX
 XX Example 1; Pages 75-77; 128pp; Japanese.
 PS
 CC The present sequence is the C. pneumoniae polypeptide antigen
 CC polypeptide A clone, 53-35. C. pneumoniae strain YK41 was
 CC cultured and genomic DNA extracted to prep. a lambda g11 DNA
 CC library. The library was then screened with an anti-YK41
 CC monoclonal antibody (Mab), which was prepd. by fusing spleen cells
 CC from a mouse infected with YK41 with myeloma P3/NS1/1- γ g4-1 to
 CC produce a Mab expressing hybridoma. The DNA cdt. was then fused
 CC with the expression vector pMDA31 to give pcPN533alpha. The
 CC plasmid was used to transform an E. coli host, which was cultured
 CC to give the antigenic polypeptide, polypeptide A. Polypeptide A
 CC and primers and probes derived from its DNA can be used in assays
 CC for the detection of polypeptide A antibodies and DNA,
 CC respectively, useful in the diagnosis of C. pneumoniae infection.
 XX
 XX Sequence 259 AA;
 SQ

Query Match 52.2%; Score 259; DB 17; Length 259;
 Best Local Similarity 100.0%; Pred. No. 2,4e-222;
 Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MSISSSGPDNOKNIMSOVLSTPGVPOODKLSGNETKOIQOTROGKNTMESDATTAG 63
 DB 1 MSISSSGPDNOKNIMSOVLSTPGVPOODKLSGNETKOIQOTROGKNTMESDATTAG 60
 QY 64 ASGKDKTSTTETETAPQGVAAAGKSSBSOKAGADTGVSGAAATTAATATKIAMOTSI 123
 DB 61 ASGKDKTSTTETETAPQGVAAAGKSSBSOKAGADTGVSGAAATTAATATKIAMOTSI 120
 QY 124 EBAKSMESTLESLSAQAQKEVEAVVVAALSGKSSGSAKLETPELPKPGVTPRSEVI 183
 DB 121 EBAKSMESTLESLSAQAQKEVEAVVVAALSGKSSGSAKLETPELPKPGVTPRSEVI 180
 QY 184 EIGLALAKAIOTLGEATKTSALSNVASTQAQADQTNKLGLEKQAIKIDKEREYOEKMAAE 243
 DB 181 EIGLALAKAIOTLGEATKTSALSNVASTQAQADQTNKLGLEKQAIKIDKEREYOEKMAAE 240
 QY 244 OKSKDLEGTMDTVNTVMIA 262
 DB 241 OKSKDLEGTMDTVNTVMIA 259

RESULT 6
 AAW01743
 ID AAW01743 standard; Protein; 259 AA.
 AC AAW01743;
 XX
 DT 22-APR-1997 (first entry)
 XX
 DE C. pneumoniae 53 kDa antigen.
 XX

KW antigen; antibody; detection; determination; epitope.
 XX Chlamydia pneumoniae.
 OS
 XX JP08304404-A.
 XX
 XX 22-NOV-1996.
 PD
 XX 28-APR-1995; 95JP-0106014.
 XX 28-APR-1995; 95JP-0106014.
 XX 28-APR-1995; 95JP-0106014.
 XX (HITB) HITACHI CHEM CO LTD.
 PA WPI; 1997-056178/06.
 DR N-PSDB; AAT59311.
 XX
 XX Detection and determination of anti-Chlamydia pneumoniae antibody -
 PT using the polypeptide C as the antigen
 PT
 XX
 XX Example 3; Page 17-18; 18pp; Japanese.
 PS
 CC The protein is that of the Chlamydia pneumoniae 53 kDa antigen. A
 CC method for the detection and determination of anti-C. pneumoniae
 CC antibodies in a sample comprises using at least 5 consecutive amino
 CC acids of the polypeptide C 73 kDa antigen (AAW01742).
 CC
 XX
 XX Sequence 259 AA;
 SQ

Query Match 52.2%; Score 259; DB 18; Length 259;
 Best Local Similarity 100.0%; Pred. No. 2,4e-222;
 Matches 259; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 4 MSISSSGPDNOKNIMSOVLSTPGVPOODKLSGNETKOIQOTROGKNTMESDATTAG 63
 DB 1 MSISSSGPDNOKNIMSOVLSTPGVPOODKLSGNETKOIQOTROGKNTMESDATTAG 60
 QY 64 ASGKDKTSTTETETAPQGVAAAGKSSBSOKAGADTGVSGAAATTAATATKIAMOTSI 123
 DB 61 ASGKDKTSTTETETAPQGVAAAGKSSBSOKAGADTGVSGAAATTAATATKIAMOTSI 120
 QY 124 EBAKSMESTLESLSAQAQKEVEAVVVAALSGKSSGSAKLETPELPKPGVTPRSEVI 183
 DB 121 EBAKSMESTLESLSAQAQKEVEAVVVAALSGKSSGSAKLETPELPKPGVTPRSEVI 180
 QY 184 EIGLALAKAIOTLGEATKTSALSNVASTQAQADQTNKLGLEKQAIKIDKEREYOEKMAAE 243
 DB 181 EIGLALAKAIOTLGEATKTSALSNVASTQAQADQTNKLGLEKQAIKIDKEREYOEKMAAE 240
 QY 244 OKSKDLEGTMDTVNTVMIA 262
 DB 241 OKSKDLEGTMDTVNTVMIA 259

RESULT 7
 AAR94580
 ID AAR94580 standard; Protein; 271 AA.
 AC AAR94580;
 XX
 DT 07-NOV-1996 (first entry)
 XX
 XX C. pneumoniae polypeptide antigen (polypeptide A) variant.
 DE
 XX Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;
 KW pcPN533alpha; primer; assay; detection; antibody; diagnosis;
 XX infection; variant.
 XX Chlamydia pneumoniae.
 OS
 XX W09609320-A1.
 XX
 XX 28-MAR-1996.
 PD

Matches 486; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 4 MSISSSGPDNKNIMSOVLSTPGCVPOODKLSGNETKQIOOTROGKNTMESDATTAG 63
Db 162 MSISSSGPDNKNIMSOVLSTPGCVPOODKLSGNETKQIOOTROGKNTMESDATTAG 221
OY 64 ASGDKTSTTETETAPQGVNAGKESSESOKAGADTGVSGAAATTAATATKIMQTSI 123
Db 222 ASGDKTSTTETETAPQGVNAGKESSESOKAGADTGVSGAAATTAATATKIMQTSI 281
OY 124 EBAASKMESTLESLOSLSNAQKKEVEAVVVAALSGKSSGSASAKLETPPELPKPGVTPRSEVI 183
Db 282 EBAASKMESTLESLOSLSNAQKKEVEAVVVAALSGKSSGSASAKLETPPELPKPGVTPRSEVI 341
OY 184 ETGLALAKAIOTLGEATKSALSNVASTOAOADOTNKLIGEKAIKIDEREYOEKMAE 243
Db 342 ETGLALAKAIOTLGEATKSALSNVASTOAOADOTNKLIGEKAIKIDEREYOEKMAE 401
OY 244 QSKDLEGTMDTVNTVMTAVSAIVISIVAAIFTCGAGLAGIAGAAGAAGAGA 303
Db 402 QSKDLEGTMDTVNTVMTAVSAIVISIVAAIFTCGAGLAGIAGAAGAAGAGA 461
OY 304 AAATTVATQITVQAVVQAVKQAVITAVRQAITAIIKAIVKSGIKAFITLVKAIAKAIK 363
Db 462 AAATTVATQITVQAVVQAVKQAVITAVRQAITAIIKAIVKSGIKAFITLVKAIAKAIK 521
OY 364 GISKVFAGTQWIAKNPKLSKVISLTSKWTGVGVVVAAPALGKGMQJSEMCOON 423
Db 522 GISKVFAGTQWIAKNPKLSKVISLTSKWTGVGVVVAAPALGKGMQJSEMCOON 581
OY 424 VAOFOKEVGILOAADMISMFTQFMQOASKIASKOTGESNEMTOKATKLGAIILKAYAI 483
Db 582 VAOFOKEVGILOAADMISMFTQFMQOASKIASKOTGESNEMTOKATKLGAIILKAYAI 641
OY 484 SGAIAGA 490
Db 642 SGAIAGA 648

```

RESULT 4
AAR94579
ID AAR94579 standard; Protein: 488 AA.
AC AAR94579;
XX
DT 07-NOV-1996 (first entry)
XX
DE Chlamydia pneumoniae polypeptide antigen (polypeptide A).
XX
KW Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;
KW PCPN533alpha; primer; assay; detection; antibody; diagnosis;
KW Infection.
XX
OS Chlamydia pneumoniae.
XX
FH Key Location/Qualifiers
FT Peptide 1.259 /note="claimed N-terminal fragment"
ET
XX
PN WO9609320-A1.
XX
PD 28-MAR-1996.
XX
PF 20-SEP-1995; 95WO-JP01896.
XX
PR 28-APR-1995; 95JP-0106011.
PR 20-SEP-1994; 94JP-0224711.
PR 28-APR-1995; 95JP-0106006.
PR 28-APR-1995; 95JP-0106008.
PR 28-APR-1995; 95JP-0106009.
PR 28-APR-1995; 95JP-0106010.
XX
PA (HITB) HITACHI CHEM CO LTD.

XX
PI Izutsu H, Matsumoto A, Obara K;
XX
DR WPI: 1996-188399/19.
DR N-PSDB: AAT14612.
XX
PT Recombinant Chlamydia pneumoniae antigen and antibodies to it
PT used for detection and assay of C. pneumoniae e.g. in clinical
PT diagnosis
XX
PS Claim 1: Pages 60-64; 128pp; Japanese.
XX
CC The present sequence is the C. pneumoniae polypeptide antigen,
CC polypeptide A. C. pneumoniae strain YK41 was cultured and genomic
CC DNA extracted to prep. a lambda gtl1 DNA library. The library was
CC then screened with an anti-YK41 monoclonal antibody (MAb), which
CC was prep. by fusing spleen cells from a mouse infected with YK41
CC with myeloma P3/NS1/1-Ag4-1 to produce a MAb expressing hybridoma.
CC The DNA obt. was then fused with the expression vector pDA431 to
CC give pcPN533alpha. The plasmid was used to transform an E. coli
CC host, which was cultured to give the antigenic polypeptide,
CC polypeptide A. Polypeptide A and primers and probes derived from
CC its DNA can be used in assays for the detection of polypeptide A
CC antibodies and DNA, respectively, useful in the diagnosis of
CC C. pneumoniae infection.
XX
SQ Sequence 488 AA:
XX
Query Match 57.5%; Score 285; DB 17; Length 488;
Best Local Similarity 100.0%; Pred. No. 2.8e-245;
Matches 285; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

OY 4 MSISSSGPDNKNIMSOVLSTPGCVPOODKLSGNETKQIOOTROGKNTMESDATTAG 63
Db 1 MSISSSGPDNKNIMSOVLSTPGCVPOODKLSGNETKQIOOTROGKNTMESDATTAG 60
OY 64 ASGDKTSTTETETAPQGVNAGKESSESOKAGADTGVSGAAATTAATATKIMQTSI 123
Db 61 ASGDKTSTTETETAPQGVNAGKESSESOKAGADTGVSGAAATTAATATKIMQTSI 120
OY 124 EBAASKMESTLESLOSLSNAQKKEVEAVVVAALSGKSSGSASAKLETPPELPKPGVTPRSEVI 183
Db 121 EBAASKMESTLESLOSLSNAQKKEVEAVVVAALSGKSSGSASAKLETPPELPKPGVTPRSEVI 180
OY 184 ETGLALAKAIOTLGEATKSALSNVASTOAOADOTNKLIGEKAIKIDEREYOEKMAE 243
Db 181 ETGLALAKAIOTLGEATKSALSNVASTOAOADOTNKLIGEKAIKIDEREYOEKMAE 240
OY 244 QSKDLEGTMDTVNTVMTAVSAIVISIVAAIFTCGAGLAGIAGA 288
Db 241 QSKDLEGTMDTVNTVMTAVSAIVISIVAAIFTCGAGLAGIAGA 265

```

RESULT 5
AAR94586
ID AAR94586 standard; Protein: 259 AA.
AC AAR94586;
XX
DT 11-NOV-1996 (first entry)
XX
DE C. pneumoniae polypeptide antigen (polypeptide A) clone 53-3S.
XX
KW Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;
KW PCPN533alpha; primer; assay; detection; antibody; diagnosis;
KW Infection; clone.
XX
OS Chlamydia pneumoniae.
XX
PN WO9609320-A1.
XX
PD 28-MAR-1996.
XX

KW vaccine; neutralising epitope.
 XX Chlamydia pneumoniae.
 XX MO9927105-A2.
 XX 03-JUN-1999. ✓
 XX 20-NOV-1998; 98WO-IB01890.
 XX 04-NOV-1998; 98US-0107078.
 XX 21-NOV-1997; 97FIR-0014673.
 XX (BEST) GENSET.
 XX Griffais R;
 XX WPI: 1999-357842/30.
 XX
 XX Genome sequence of Chlamydia pneumoniae
 XX
 XX Page 1223-1224; Disclosure: 1912pp; English.

CC AAY34584-Y35879 represent the proteins encoded by all the open reading
 CC frames in the complete genome (see AAY31990) of Chlamydia pneumoniae.
 CC C. pneumoniae causes respiratory disease such as pneumonia and
 CC bronchitis and is thought to be a contributing factor in heart
 CC disease, sarcoidosis, sinusitis, purulent otitis media, erythema
 CC nodosum or pharyngitis. The polypeptides encoded by the open reading
 CC frames of the C. pneumoniae genome (see AAY34584-Y35879) can be used in
 CC immunogenic compositions as vaccines. Vectors containing C. pneumoniae
 CC nucleotide sequences can also be used as immunogenic compositions,
 CC especially where the vector directs the expression of a neutralising
 CC epitope of C. pneumoniae.

XX Sequence 496 AA:

Query Match 98.2%; Score 487; DB 20; Length 496;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 DTNNSTSSSSGPNOKNIMQVLTSTPGVPPOODKLSGNETKIOOTROCKNEMESD 60
 DB 1 DTNNSTSSSSGPNOKNIMQVLTSTPGVPPOODKLSGNETKIOOTROCKNEMESD 60
 QY 61 IAGASGDKTSTTKEFTAPQOGVAAGKESSESOKAGADTGVSAAATTAATNATKIAMQ 120
 DB 61 IAGASGDKTSTTKEFTAPQOGVAAGKESSESOKAGADTGVSAAATTAATNATKIAMQ 120
 QY 121 TSIEEASKSNESTLESLSQISAQOMEVEAVVAALSGSSGSAKLETPPELPGVTPRS 180
 DB 121 TSIEEASKSNESTLESLSQISAQOMEVEAVVAALSGSSGSAKLETPPELPGVTPRS 180
 QY 181 EYEIIGLALAKAQTOLGEATKLSALSNVASTQADOTNKLGLKQAIKIDKEEEYOEAK 240
 DB 181 EYEIIGLALAKAQTOLGEATKLSALSNVASTQADOTNKLGLKQAIKIDKEEEYOEAK 240
 QY 241 AABOKSKDELTGMDVTNVAIAVSAITVISAALFTGAGLAGIAAAGAAAGAAAGAA 300
 DB 241 AABOKSKDELTGMDVTNVAIAVSAITVISAALFTGAGLAGIAAAGAAAGAAAGAA 300
 QY 301 AGAAAAATTAQTITVOAVVQAVVAVITAVROAITAIAKAAVSGIKAFIKTLVAIAKA 360
 DB 301 AGAAAAATTAQTITVOAVVQAVVAVITAVROAITAIAKAAVSGIKAFIKTLVAIAKA 360
 QY 361 ISGIGISVAKGOMIAKAPKLSKVISLTSKWTVYGVVVAAPALGKIMQOLSEM 420
 DB 361 ISGIGISVAKGOMIAKAPKLSKVISLTSKWTVYGVVVAAPALGKIMQOLSEM 420
 QY 421 QONVAQFOKEVQKLOAAADIMSMFTQFMOQASKIASKOTGESNEMTOKRTXGAGIILAY 480
 DB 421 QONVAQFOKEVQKLOAAADIMSMFTQFMOQASKIASKOTGESNEMTOKRTXGAGIILAY 480

QY 481 AATGAI 487
 DB 481 AATGAI 487

RESULT 3
 AAR94584

ID AAR94584 standard; Protein; 649 AA.

AC AAR94584;

DE 11-NOV-1996 (first entry)

XX DHFR/C. pneumoniae antigen fusion protein (polypeptide B).

XX Polypeptide antigen; polypeptide B; strain YK41; plasmid; probe;

XX PCPN533T; primer; assay; detection; antibody; diagnosis;

XX Chlamydia pneumoniae.

XX Synthetic.

FT Key Location/Qualifiers

FT Protein 1..160 /label= dihydrofolate reductase

FT Protein 162..649 /label= C. pneumoniae antigen

PN MO9609320-A1.

PD 28-MAR-1996.

PF 20-SEP-1995; 95WO-JP01896.

PR 28-APR-1995; 95JP-0106011.

PR 20-SEP-1994; 94JP-0224711.

PR 28-APR-1995; 95JP-0106006.

PR 28-APR-1995; 95JP-0106008.

PR 28-APR-1995; 95JP-0106009.

PR 28-APR-1995; 95JP-0106010.

PA (HITB) HITACHI CHEM CO LTD.

PI Izutsu H, Matsumoto A, Obara K;

DR WPI: 1996-188399/19.

DR N-PSDB: AAT14618.

PT Recombinant Chlamydia pneumoniae antigen and antibodies to it

PT used for detection and assay of C. pneumoniae e.g. in clinical

PT diagnosis

PS Claim 22; Pages 97-103; 128pp; Japanese.

XX The present sequence is the dihydrofolate reductase (DHFR)/

XX C. pneumoniae polypeptide antigen fusion protein, polypeptide B.

XX C. pneumoniae strain YK41 was cultured and genomic DNA extracted

XX with an anti-YK41 monoclonal antibody (MAb), which was then prep. by

XX fusing spleen cells from a mouse infected with YK41 with myeloma

XX P3/MS1/1-Ag4-1 to produce a MAb expressing hybridoma. The DNA

XX obd. was then fused with DHFR DNA and the expression vector

XX pADNA43 to give PCPN533T. The plasmid was used to transform an

XX E. coli host, which was cultured to give the antigenic polypeptide

XX fusion protein, polypeptide B. Polypeptide B and primers and

XX detection of antigenic polypeptide antibodies and DNA, respectively,

XX useful in the diagnosis of C. pneumoniae infection.

Sequence 649 AA;

Query Match 77.8%; Score 386; DB 17; Length 649;
 Best Local Similarity 99.8%; Pred. No. 0;

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 16:32:30 ; Search time 31 Seconds
(without alignments)
2132.010 Million cell updates/sec

Title: US-09-889-314-2

Perfect score: 496
Sequence: 1 DTNMSISSSGPDKNINMS.....LKAYAAISGAIGAKHTNMF 496

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 908470 seqs, 133250620 residues

Word size: 0

Total number of hits satisfying chosen parameters: 908470

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

A.Geneseq_101002.*
1: /SID2/gcgdata/geneseq/geneseq-emb1/AA1980.DAT:*
2: /SID2/gcgdata/geneseq/geneseq-emb1/AA1981.DAT:*
3: /SID2/gcgdata/geneseq/geneseq-emb1/AA1982.DAT:*
4: /SID2/gcgdata/geneseq/geneseq-emb1/AA1983.DAT:*
5: /SID2/gcgdata/geneseq/geneseq-emb1/AA1984.DAT:*
6: /SID2/gcgdata/geneseq/geneseq-emb1/AA1985.DAT:*
7: /SID2/gcgdata/geneseq/geneseq-emb1/AA1986.DAT:*
8: /SID2/gcgdata/geneseq/geneseq-emb1/AA1987.DAT:*
9: /SID2/gcgdata/geneseq/geneseq-emb1/AA1988.DAT:*
10: /SID2/gcgdata/geneseq/geneseq-emb1/AA1989.DAT:*
11: /SID2/gcgdata/geneseq/geneseq-emb1/AA1990.DAT:*
12: /SID2/gcgdata/geneseq/geneseq-emb1/AA1991.DAT:*
13: /SID2/gcgdata/geneseq/geneseq-emb1/AA1992.DAT:*
14: /SID2/gcgdata/geneseq/geneseq-emb1/AA1993.DAT:*
15: /SID2/gcgdata/geneseq/geneseq-emb1/AA1994.DAT:*
16: /SID2/gcgdata/geneseq/geneseq-emb1/AA1995.DAT:*
17: /SID2/gcgdata/geneseq/geneseq-emb1/AA1996.DAT:*
18: /SID2/gcgdata/geneseq/geneseq-emb1/AA1997.DAT:*
19: /SID2/gcgdata/geneseq/geneseq-emb1/AA1998.DAT:*
20: /SID2/gcgdata/geneseq/geneseq-emb1/AA1999.DAT:*
21: /SID2/gcgdata/geneseq/geneseq-emb1/AA2000.DAT:*
22: /SID2/gcgdata/geneseq/geneseq-emb1/AA2001.DAT:*
23: /SID2/gcgdata/geneseq/geneseq-emb1/AA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	496	100.0	496	21	AA083322
2	487	98.2	496	20	AAV35449
3	386	77.8	649	17	AA094584
4	285	57.5	488	17	AA094579
5	259	52.2	259	17	AA094586
6	259	52.2	259	18	AA01743
7	259	52.2	271	17	AA094580
8	259	52.2	302	21	AA083323
9	259	52.2	432	17	AA094585
10	14	2.8	14	21	AA083335

11	14	2.8	14	21	AA083336	Epitope derived fr
12	11	2.2	1937	22	AB058985	Drosophila melanog
13	10	2.0	129	23	AB031472	Human ORF45, prote
14	10	2.0	720	23	AA014994	Laminin-related pr
15	10	2.0	755	22	AA050055	Propionibacterium
16	9	1.8	9	21	AA083324	Epitope derived fr
17	9	1.8	9	21	AA083326	Epitope derived fr
18	9	1.8	9	21	AA083330	Epitope derived fr
19	9	1.8	9	21	AA083333	Epitope derived fr
20	9	1.8	9	21	AA083334	Epitope derived fr
21	9	1.8	111	16	AA072659	Cladosporium herba
22	9	1.8	148	21	AA011695	Arabidopsis thalia
23	9	1.8	149	21	AA011694	Arabidopsis thalia
24	9	1.8	153	22	AB038670	Peptide #6176 enco
25	9	1.8	153	22	AB023746	Protein #5745 enco
26	9	1.8	153	22	AA059299	Human brain expres
27	9	1.8	153	22	AA071846	Human bone marrow
28	9	1.8	153	22	AA019335	Peptide #5769 enco
29	9	1.8	153	22	AA032131	Peptide #6168 enco
30	9	1.8	153	23	AB041659	Human peptide enco
31	9	1.8	159	21	AA011693	Arabidopsis thalia
32	9	1.8	169	20	AA042438	Amno acid sequenc
33	9	1.8	169	21	AA011696	Human peripheral-t
34	9	1.8	169	21	AA093604	Protein encoded by
35	9	1.8	172	21	AA038041	Fragment of human
36	9	1.8	354	22	AB005860	Novel human diago
37	9	1.8	379	21	AA010364	Arabidopsis thalia
38	9	1.8	382	21	AA010363	Arabidopsis thalia
39	9	1.8	477	23	AB092256	Herbicideally activ
40	9	1.8	481	20	AA037632	Protein which is s
41	9	1.8	513	22	AA078562	Human protein SEQ
42	9	1.8	518	22	AA079546	Human protein SEQ
43	9	1.8	555	21	AA012164	Hydrophobic domain
44	9	1.8	555	22	AA073670	Human oxidoreducta
45	9	1.8	696	22	AB058615	Drosophila melanog
46	9	1.8	1001	22	AB058298	Drosophila melanog
47	9	1.8	1024	22	AB061798	Drosophila melanog
48	9	1.8	1615	22	AA081211	Mycobacterium tube
49	8	1.6	8	21	AA083328	Epitope derived fr
50	8	1.6	27	21	AA051199	Human SCA7 protein
51	8	1.6	28	21	AA059113	N. clavipes spider
52	8	1.6	30	21	AA059124	N. clavipes spider
53	8	1.6	33	19	AA053349	Nephila clavipes s
54	8	1.6	33	21	AA059073	N. clavipes spider
55	8	1.6	39	22	AB030746	Peptide #3397 enco
56	8	1.6	39	22	AB035923	Peptide #3429 enco
57	8	1.6	39	22	AB021332	Peptide #3331 enco
58	8	1.6	39	22	AA056726	Human brain expres
59	8	1.6	39	22	AA069104	Human bone marrow
60	8	1.6	39	22	AA016940	Peptide #3374 enco
61	8	1.6	39	22	AA029425	Peptide #3462 enco
62	8	1.6	39	22	AA004633	Peptide #3315 enco
63	8	1.6	39	22	AB021332	Human peptide enco
64	8	1.6	81	22	AB025327	Novel human diago
65	8	1.6	84	18	AA019263	Lactobacillus amyl
66	8	1.6	109	22	AA093342	Human polypeptide,
67	8	1.6	111	17	AA006538	C. herbarum allerg
68	8	1.6	122	21	AA043847	Zea mays protein f
69	8	1.6	122	21	AA043950	Zea mays protein f
70	8	1.6	129	21	AA051200	Human SCA7 protein
71	8	1.6	138	21	AA043946	Zea mays protein f
72	8	1.6	138	21	AA043949	Zea mays protein f
73	8	1.6	164	20	AA044018	Human glutaredoxin
74	8	1.6	164	21	AA012767	Human glutaredoxin
75	8	1.6	165	21	AA043945	Zea mays protein f
76	8	1.6	165	21	AA043948	Zea mays protein f
77	8	1.6	179	21	AA024334	Arabidopsis thalia
78	8	1.6	179	21	AA046905	Arabidopsis thalia
79	8	1.6	197	23	AA017565	M. catenalis WCAL
80	8	1.6	209	22	AB011564	Human secreted pro
81	8	1.6	211	21	AA040499	Arabidopsis thalia
82	8	1.6	213	22	AB059343	Drosophila melanog
83	8	1.6	221	22	AA092753	C. glutamicum prote


```

Db 1058 AAAAAGCTGACGAGCGCGCAAAATTAATTCATTCGCTGACCGCGCTGACCGCGCTAT 1117
OY 103 SerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSer 122
Db 1118 GCACAGCTGAAGCC-----GGG 1135
OY 123 ILeGluAlaAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerAla 142
Db 1136 GTAGACAGCGCGGAAAGACGACGACGACGACGACGACGACGACGACGACGACGACGAT 1195
OY 143 AlaGlnMetLys-----GluValGluAlaValAlaValAlaAlaLeuSer 157
Db 1196 GCGACGCTTAAGACGACGACGACGACGACGACGACGACGACGACGACGATTCG 1255
OY 158 GlyLysSerSerGlySerAlaLysLeuGluThrProGluLeuProLysProGlyValThr 177
Db 1256 ACCAAATTCAGGGAACGCTATGCCGCTTCAGAAATCAGCTTCCAGGGGTGACAG 1315
OY 178 ProArg---SerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeu 196
Db 1316 GATATCTGTCAATATCGCGCGCTCCTACTATGCTCATGCGCATGTTATTGAGATTGTG 1375
OY 197 GlyGluAlaThrLysSerAlaLeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGln 216
Db 1376 GGCAGAAATATACGAGAAAGACCTGCAAAACGATCTTGCGCTT----- 1417
OY 217 ThrAsnLysLeuGlyLeuGlyLysGlnAlaIleLysIleAspLysGluArgGluTyr 236
Db 1418 TTCACGCCCTTGACGAGAGAGCGCTCAGCGC---GAGATGGAGAAATCGCGCTGAATTC 1474
OY 237 GlnGluMetLysAlaIleGluGlnLysSerLysAspLeuGluGlyThrMetAspThrVal 256
Db 1475 CAGGAA-----GAGACCGCGAAGCGGAGAGAACGACCGCATTTATGGAGTATC 1525
OY 257 AsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIle 276
Db 1526 GGGAAAGTCTCGCGCGCTG-----CTAACCATTTGTCAGCGTTGCGCGCTGTT 1576
OY 277 PheThrCysGlyValAclyLeuAlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAlaAla 296
Db 1577 TTTACCGGTGGCGGAGTGTGGCGCTGCGTGGGACTTGGCTGCGTA----- 1624
OY 297 AlaGlyGlyAlaAlaGlyAlaAlaAlaThrThrValAlaThrGlnIleThrValGln 316
Db 1625 -----ATGCTGCCGATGAATT----- 1642
OY 317 AlaValAlaGlnAlaValLysGlnAlaVal---IleThrAlaValArgGlnAlaIleThr 335
Db 1643 -----GTGAAGGCGGAGCGGAGGTGCTTTATTCAGACGCGCTAAAC 1687
OY 336 AlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLys 355
Db 1688 CCGATTATGAGCATGTG-----CTGAAGCCGTTAATGAG 1723
OY 356 AlaIleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMet 375
Db 1724 CTGATTGGCAAGCGCATTACCAAGCGCTG-----GAAGATTAGCGCTC 1768
OY 376 IleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysTyrVal 395
Db 1769 GATAAGAAACGACGACGATGGCGGACGATTTGTCGATTTGCGCGCATTTGCC 1828
OY 396 ThrValGlyValGlyValVal-----AlaAlaProAlaLeuGly 409
Db 1829 ATGTGTGGCTCATTTGTGTGTCGACATTGTGCGGAAAGCGCGGCAAACTGGGT 1888
OY 410 LysGlyIleMetGlnMetGlnLeuSerGluMetGlnGlnAsnValAlaGlnPheGlnLys 429
Db 1889 AACGCGCTGAGCAAAATG-----ATGGCGAAACGATTAAGAGTTGGTGCTT 1936
OY 430 GluValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPheTyrPgl 449

```

```

Db 1937 AACGTGCTGAAA---CAGTTGGCGGCAAAACGACGACGAACTCTTACCAGGGGATGCA 1993
OY 450 Gln-----AlaSerLysIleAlaSerLysGlnThrGlyIle 461
Db 1994 CGATTACTAGCGGCTGTGGGTAAATGTGGGTAGCAGATGGCCCTGCAACGAATGCTTA 2053
OY 462 SerAsnGluMet-----ThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLys 478
Db 2054 AGTAAGACCTGTAGTAAATACCTAAATTAAGTGGCGTTGGCATGGAAGTCAACGAAT 2113
OY 479 AlaTyrAlaAlaIleSerGlyAlaIleAla 488
Db 2114 ACCGACGCCAGTCAGCGCGTGTGGTGGCC 2143

```

Search completed: January 27, 2003, 15:06:56
Job time : 3020 secs

US-09-889-314-2 (1-496) x STU25631 (1-3622)

QY 28 GlnGlyValProGlnGlnAspLysLeuSerGlyAsnGluThrLysGlnIleGlnGlnThr 47
 Db 908 CAGGCGCATATGATGACACAAAAGAGATGATTCAGATTCGAAATTCACAGACG 967

QY 48 ArgGlnGlyLysAsnThrGluMetGluSerAspAlaThrIleAlaGlyLysSerGlyLys 67
 Db 968 GCTCTGGGA-----GAGGCTCAGAGGCGGAGCATCTCTATGAAAGCCGATCAAA 1018

QY 68 AspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGlyAlaAlaGly 87
 Db 1019 -----AAGACGGATACCGCCAAAGTGTATTATGACCTCGACCC 1057

QY 88 LysGluSerSerGluSerGlnLysAla-----GlyAlaAspThrGlyVal 102
 Db 1058 AAAAACTGACGCGGCGGAAATTAATTCATTCGTCGACCGCGCTGATGACCGCTAT 1117

QY 103 SerGlyAlaAlaAlaThrThrLysSerAsnThrAlaThrLysIleAlaMetGlnThrSer 122
 Db 1118 GCACAAGCTGAAGCC-----GCG 1135

QY 123 IleGlnGluAlaSerLysSerMetGluSerThrLeuGluSerLeuSerLeuSerAla 142
 Db 1136 GTAGAACAGCGCGGAAAGACGACAGAGCGGAAAGAGCGCTTATGATTAAGCCACGAT 1195

QY 143 AlaGlnMetLys-----GluValGluAlaValAlaValAlaAlaLeuSer 157
 Db 1196 GCGACGGTTAAAGCAGCAGCAGACCCCAAGCGGAAAGCGGAAAGCGATTAATCTGTG 1255

QY 158 GlyLysSerSerGlySerAlaLysLeuGluThrProGluLeuProLysProGlyValThr 177
 Db 1256 ACCAATTCACAGGAGCGGTAATCCCGCTCGATTCAGATTCAGTTCCAGAGGTACAGAG 1315

QY 178 ProArg---SerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeu 196
 Db 1316 GATAATCTGCTCAATGTCGCCCTCCTACTATGCTCATGCGCATGTTATGAGATTTGTG 1375

QY 197 GlyAlaIleThrLysSerAlaLeuSerAsnThrLysSerThrGlnAlaGlnAlaSerAla 216
 Db 1376 GCGAAATATCAGGAAAGAGCGCTCAACATCTGCGCT----- 1417

QY 217 ThrAsnLysLeuGlyLeuGlnGlnAlaIleLysIleAspLysGluArgGluGluThr 236
 Db 1418 TTCACGCGCTTGCAGAGAGCGCTCAGCGC---GAGATGAAAGAAATCGCTGAATTC 1474

QY 237 GlnGluMetLysAlaAlaGlnGlnLysSerLysAspLeuGluGlyThrMetAspThrVal 256
 Db 1475 CAGGAA-----GAGACCGCGCAAGCGGAGGAAACGAAACCGCATTAATGATGTATC 1525

QY 257 AsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaIle 276
 Db 1526 GCGAAGATCCTCGCGCGCGCTG-----CTAACCATGTCACAGCTGTGTGCGCGCTGT 1576

QY 277 PheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaValGlyAlaAlaAla 296
 Db 1577 TTTCACCGGTGGCGAGCTGCGCTGCTGCGGTGCGATCTGCGGTA----- 1624

QY 297 AlaGlyLysAlaAlaGlyAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGln 316
 Db 1625 -----ATGCTGCCCGATGAAT----- 1642

QY 317 AlaValAlaGlnAlaValLysGlnAlaVal---IleThrAlaValArgGlnAlaIleThr 335
 Db 1643 -----GTGAAGCGCGCGCAGCGAGATGCTGTTATTCACGAGCGGCTTAAC 1687

QY 336 AlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLys 355
 Db 1688 CCGATTTATGAGCATGTG-----CTGAAGCGCGTTAATGAGG 1723

QY 356 AlaIleLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMet 375
 Db 1723 -----GlyAlaAspThrGlyVal 102

Db 1724 CTGATTGGCAAGCGCATTCACCAAGCGCTG-----GAAGATTTAGCGCTC 1768

QY 376 IleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysThrVal 395
 Db 1769 GATTAAGAAACCGCGAGAGATGGCGGAGCATTTGTGTGGATTTGCGCGCATTTGCC 1828

QY 396 ThrValGlyValGlyValAlaVal-----AlaAlaProAlaLeuGly 409
 Db 1829 ATGTGTGGCGGTATTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1888

QY 410 LysGlyIleMetGlnMetGlnLeuSerGluMetGlnGlnAsnValAlaGlnPheGlnLys 429
 Db 1889 AACGGCGTCGACCAAAATG-----ATGGCCGAACGATTAGACATTTGCTGCTT 1936

QY 430 GluValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPheThrGln 449
 Db 1937 AACGTGCTGAAA---CACTTGGCGCAAAACGCGACAAACCTCTTACCAAGGGATGCA 1993

QY 450 Gln-----AlaSerLysIleAlaSerLysGlnThrGlyLys 461
 Db 1994 CGTATTACTACGGCTGCGGTAAATGTGGTAGCAAGATGGCGCTGCAACGAATGCTTA 2053

QY 462 SerAsnGluMet-----ThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLys 478
 Db 2054 AGTAAGAGCTGTGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2113

QY 479 AlaThrAlaAlaIleSerGlyAlaIleAla 488
 Db 2114 ACCGACGCCAGTCAGCGGTGTGTGTGCC 2143

RESULT 45
 AR082441
 LOCUS AR082441 3622 bp DNA linear PAT 31-AUG-2000
 DEFINITION Sequence 7 from patent US 5972899.
 ACCESSION AR082441
 VERSION AR082441.1 GI:10009167
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 3622)
 AUTHORS Zychlinsky, A. and Chen, Y.
 TITLE Apoptosis induced by Shigella IPAB
 JOURNAL Patent: US 5972899-A 7 26-OCT-1999;
 FEATURES
 source
 1..3622
 /organism="unknown"
 BASE COUNT 1049 a 759 c 1008 g 806 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0.00316 Length: 3622
 Score: 218.50 Matches: 116
 Percent Similarity: 41.02% Conservative: 85
 Best Local Similarity: 23.67% Mismatches: 182
 Query Match: 9.22% Indels: 107
 DB: 6 Gaps: 20

US-09-889-314-2 (1-496) x AR082441 (1-3622)

QY 28 GlnGlyValProGlnGlnAspLysLeuSerGlyAsnGluThrLysGlnIleGlnGlnThr 47
 Db 908 CAGGCGCATATGATGACACAAAAGAGATGATTCAGATTCGAAATTCACAGACG 967

QY 48 ArgGlnGlyLysAsnThrGluMetGluSerAspAlaThrIleAlaGlyLysSerGlyLys 67
 Db 968 GCTCTGGGA-----GAGGCTCAGAGGCGGAGCATCTCTATGAAAGCCGATCAAA 1018

QY 68 AspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGlyAlaAlaGly 87
 Db 1019 -----AAGACGGATACCGCCAAAGTGTATTATGACCTCGACCC 1057

QY 88 LysGluSerSerGluSerGlnLysAla-----GlyAlaAspThrGlyVal 102


```

OY      479  ALATYALALAIleSergLYAlAIleAla 488
DB      2082 ACCGACGCCAGCAGCCGCTGGCTTCC 2111

RESULT 42
LOCUS   AL627276/c 274050 bp DNA linear BCT 06-JUN-2002
DEFINITION Salmonella enterica serovar Typhi (Salmonella typhi) strain CT18, complete chromosome, segment 12/20.
ACCESSION AL627276 AL513382
VERSION   AL627276.1 GI:16503805
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS 1 (bases 1 to 274050)
Parkhill,J., Dougan,G., James,K.D., Thomson,N.R., Pickard,D.,
Main,J., Churcher,C., Mungall,K.L., Bentley,S.D., Holden,M.T.G.,
Sebaihia,M., Baker,S., Basham,D., Brooks,K., Chillingworth,T.,
Conerton,P., Cronin,A., Davis,P., Davies,R.M., Dowd,L., White,N.,
Jagels,K., Krogh,A., Larsen,T.S., Leather,S., Hien,T.T., Holroyd,S.,
Parr,C., Quail,M., Rutherford,K., Simmonds,M., Skelton,J.,
Stevens,K., Whitehead,S. and Barrall,B.G.
Complete genome sequence of a multiple drug resistant Salmonella
enterica serovar Typhi CT18
Nature 413 (6858), 848-852 (2001)
JOURNAL MEDLINE 21534947
PUBMED 11677608
REFERENCE 2 (bases 1 to 274050)
Parkhill,J.
Direct Submission
Submitted (25-OCT-2001) Submitted on behalf of the Salmonella
sequencing team, Sanger Centre, Wellcome Trust Genome Campus,
Hinxton, Cambridge CB10 1SA, UK
E-mail: parkhill@sanger.ac.uk
COMMENT
Notes:
Details of S. typhi sequencing at the Sanger Centre are available
on the World Wide Web.
Location/Qualifiers
1. 274050
/organism="Salmonella enterica subsp. enterica serovar
Typhi"
/strain="CT18"
/db_xref="taxon:90370"
197..201
/note="possible RBS"
208..897
/gene="STY2840"
/note="ung"
208..897
/gene="STY2840"
/EC_number="3.2.2.-"
/note="Orthologue of E. coli ung (UNG_ECOLI); Fasta hit to
UNG_ECOLI (228 aa), 91% identity in 228 aa overlap"
/codon_start=1
/transl_table=11
/product="uracil-DNA glycosylase"
/protein_id="CAD05831.1"
/db_xref="GI:16503806"
/db_xref="SPRMBL:08XGK6"
/translation="MATELTMDVLDKDEKQPFYINTLHTVAGEKSGITVYPPQKY
FNAFRTLEGDVAVVILGDPYHGPQAGHLASVVRGTLAPPSLVNMYKELASLPG
FVAPAHYTESAKAGVLLNTLVLTIRACQASHASLIGWTFIDKVI SLINQRECVI
FLMGSHAOKKAIGIDPQRHHLIKAPHPSLSAHRGFFGCGNHFALTNQMLEGCEKTI
DMTPVLPASEE"
322..867
/gene="STY2840"
/note="pfam match to entry PF00315 UNG, Uracil-DNA
glycosylase, score 451.20, E-value 9e-132"

misc_feature
376..405
/gene="STY2840"
/note="PS00130 Uracil-DNA glycosylase signature"
complement(1013..2050)
/gene="STY2841"
/note="yif"
complement(1013..2050)
/gene="STY2841"
/note="Orthologue of E. coli yif (YIF_ECOLI); Fasta hit
to YIF_ECOLI (345 aa), 88% identity in 345 aa overlap"
/codon_start=1
/transl_table=11
/product="putative RNA methyltransferase"
/protein_id="CAD05832.1"
/db_xref="GI:16503807"
/db_xref="SPRMBL:08XFT1"
/translation="MDELKNKSGKVVVYVRSDSDSKRTNHPRTGKGGRPKSRT
DGGRRPARDERNNQSRDRKHETSPWRTVSAPGDETPEKVDGICSGKSPIDEVLR
ORAEFRVYGCENCOALFOSRPDAIVRAMPYQSVTRFRKALRMANKRAHYVDEA
ELAKASSTEHNGVCFLTKRNGTYKQWYKQADDCVLAEDVANPNLGGMMRSC
AHRCVKGVVQDVALLESGAIVTAEGAGAHVPIGESIVDLDFRQAGTVVTS
SDRQALFSTLPEKMWLVIGREYDLPFAARBPDDLVCVINGTGVESLNVATGV
LLEWWRONKA"
complement(1040..1462)
/gene="STY2841"
/note="pfam match to entry PF00588 SpoD_methylase, Spo
RNA methylase family, score 181.50, E-value 1.4e-50"
2244..2248
/note="possible RBS"
2254..2673
/gene="STY2842"
/note="trxC; yfiG"
2254..2673
/gene="STY2842"
/note="Orthologue of E. coli yfiG (YHI2_ECOLI); Fasta hit
to YHI2_ECOLI (139 aa), 94% identity in 139 aa overlap"
/codon_start=1
/transl_table=11
/product="thioredoxin 2"
/protein_id="CAD05833.1"
/db_xref="GI:16503808"
/db_xref="SPRMBL:08ZAJ6"
/translation="MNTVCTHCOAINTIRIGDRILODAKCGRGHLPDGEVYNAGET
GIRSIPIMIAIFKRGVYDMLNGAVPAPPSWLNEL"
2353..2670
/gene="STY2842"
/note="PS00194 Thioredoxin family active site"
2746..3426
/gene="STY2843"
2746..3426
/gene="STY2843"
/note="Orthologue of E. coli YIFP_ECOLI; Fasta hit to
YIFP_ECOLI (232 aa), 88% identity in 226 aa overlap"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="CAD05834.1"
/db_xref="GI:16503809"
/db_xref="SPRMBL:08ZAJ5"
/translation="MTNNAVQLRAERLARATRPPLANGNRVRCORCLLPKSLCD
TLPSQAKRECLVMEDEPMKPSWNGRLIADILPDTAFOMSPTREPOALDELVQHP
DVPQIVYFASVAGEAREVTSPPAGKPLFTMLGDTPEARKMRKSPYLDHLPVIS
VLSIRLSAARLEHIAEGCYCAFAVALILLDAGPEAATSLGEHFIKRTYIAGKI
OHPGNVTA"
3480..6140
/gene="STY2844"
3480..6140
/gene="STY2844"

```

QY 410 LysGIYIleMeGIuMeGIuLeuSerGIuMeGIuGIuGIuAsnValAlaGlnPheGlnLys 429
Db 1857 AACGGCGTGAAGAAATG-----ATGGCGCAACCATTAACAGTGGTGCT 1904
QY 430 GluValGIYLYsLeuGlnAlaAlaAspMetIleSerMetPheThrGlnPheTrpGln 449
Db 1905 AACGGCTGAAA---CAGTTGGCACAAACGCGACCAAACTCTTACCAGGGGATGCA 1961
QY 450 Gln-----AlaSerLYsIleAlaSerLYsGlnThrGln 461
Db 1962 CGTATTACTAGCGCGCTGGTAATGTGGTACAGATGGCGCTCAACGAATGCTTA 2021
QY 462 SerAsnGluMet-----ThrGlnLYsAlaThrLYsGlnAlaGlnIleLeuLys 478
Db 2022 ACTAAGAAGCTGTGATGTAATACCTTAATAAATAGTGGCTTGCGATGTAATCAAGAT 2081
QY 479 AlaTrpAlaAlaIleSerGIYAlaIleAla 488
Db 2082 ACCGACGCCACTCAGCGGTGGGTGCC 2111

RESULT 41
AR082442 5393 bp DNA linear PAT 31-AUG-2000
LOCUS AR082442 Sequence 9 from patent us 5972899.
DEFINITION AR082442
ACCESSION AR082442 GI:10009168
VERSION AR082442.1 GI:10009168
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5393)
AUTHORS Zychlinsky, A. and Chen, Y.
TITLE Apoptosis induced by Shigella IPaB
JOURNAL Patent: US 5972899-A 9/26-Oct-1999;
FEATURES
source 1. 5393
Location/Qualifiers
/organism="unknown"

BASE COUNT 1584 a 1156 c 1458 g 1195 t
ORIGIN

Alignment Scores:
Pred. No.: 0.00397 Length: 5393
Score: 220.50 Matches: 116
Percent Similarity: 41.02% Conservative: 85
Best Local Similarity: 23.67% Mismatches: 182
Query Match: 9.31% Indels: 107
DB: 6 Gaps: 20

US-09-889-314-2 (1-496) x AR082442 (1-5393)

QY 28 GlnGIYValProGlnGlnAspLYsLeuSerGIYAsnGluThrLYsGlnIleGlnGlnThr 47
Db 876 CAGCGGATGATGATGATCAAAAGAGATGGGATTCAGATTCGAAAGAAATTCAGACG 935
QY 48 ArgGlnGIYLYsAsnThrGluMetGluSerAspAlaThrIleAlaGlnLYsLeuSerLYs 67
Db 936 GCTGTGGG-----GAGGCTCAGAGGCGACGATCTCTATGAAACCGACATCAAA 986
QY 68 AspLYsThrSerThrThrLYsThrGlnThrAlaProGlnGlnGlnLYsAlaIleGln 87
Db 987 -----AAGACGATACCGCCAGACGCTTATGACCTGCGGCC 1025
QY 88 LysGluSerSerGluSerGlnLYsAla-----GlyAlaAspThrGlnLYs 102
Db 1026 AAAAAATCAGCAGCGCAAAATTAATTCATGCTGACCCAGCTGACCCGCGCTAT 1085
QY 103 SerGlnLYsAlaAlaAlaThrThrAlaSerAsnThrAlaThrLYsIleAlaMetGlnThrSer 122
Db 1086 GCACAGCTGAAGC-----GGC 1103
QY 123 IleGlnGlnAlaSerLYsSerMetGluSerThrLeuGluSerLeuGlnSerLeuSerAla 142
Db -----GGC 1103

Db 1104 GTAGAACAGCGCGGAAAGAACGACAGCGGAAAGAGCGCTTATGATTAAGCCACGAT 1163
QY 143 AlaGlnMetLYs-----GluValGlnAlaValAlaValAlaLeuSer 157
Db 1164 GCGACGTTAAAGCGACGACAGCCCAAGCGAAAGCCGAGAAAGCGGATTAATTCG 1223
QY 158 GlyLYsSerSerLYsSerAlaLYsLeuGlnThrProGluLeuProLYsProGlnLYsThr 177
Db 1224 ACCAATTCACAGGAAGGCTAATGCCCTCTCAGATTCAGATTCACAGGTCAGCAG 1283
QY 178 ProArg---SerGlnValIleGlnIleGlnLeuAlaLeuAlaLYsAlaIleGlnThrLeu 196
Db 1284 GATTAATCTCAAAATGTGCGCCGCTCACTATGCTCAGATGGCCATGTTATGAGTTGTG 1343
QY 197 GlyGlnAlaThrLYsSerAlaLeuSerAsnTrpAlaSerThrGlnAlaGlnAlaAspGln 216
Db 1344 GGCAAAATATCGGAAGAAAGCTGCAAAACGATCTTGCGCTT----- 1385
QY 217 ThrAsnLYsLeuGlnLYsGlnAlaIleLYsIleAspLYsGluArgGlnGluTrp 236
Db 1386 TTCACGCGCTTCACAGAAAGCGGTCAGGCG---GAGATGGAAGAAATCGGCTGAATTC 1442
QY 237 GlnGluMetLYsAlaAlaGlnGlnLYsSerLYsAspLeuGlnLYsThMetAspThrVal 256
Db 1443 CAGGAA-----GAGACGCGCAAGCCGAGAAACGACATTATGAGATGTATC 1493
QY 257 AsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIle 276
Db 1494 GGGAAAGTCTCGCGCGCTG-----CTAACATTGTAGCGTGTGCGCTGCT 1544
QY 277 PheThrCYsGlnLYsGlnLYsLeuAlaIleGlnLYsLeuAlaIleValAlaValAlaIle 296
Db 1545 TTTACCGGTGGGCGCAGTGTGCGCTGCGGTGGCGATTCGCGGTA----- 1592
QY 297 AlaGlnLYsAlaAlaGlnLYsAlaAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGln 316
Db 1593 -----ATNGGCGCCGAGAAATY----- 1610
QY 317 AlaValAlaGlnAlaValLYsGlnAlaVal---IleThrAlaValArgGlnAlaIleThr 335
Db 1611 -----GTGAAGCGGCGACGCGGCGTCTGTTATTCAGACGCGCTAAAC 1655
QY 336 AlaAlaIleLYsAlaAlaValLYsSerGlnLYsAlaPheIleLYsThrLeuValLYs 355
Db 1656 CCGATTATGACACATGT-----CTAACCGCTTATTCGAG 1691
QY 356 AlaIleAlaLYsAlaIleSerLYsGlnLYsSerLYsValPheAlaLYsGlnLYsThrGlnMet 375
Db 1692 CTGATTGGCAAGCGATTAACCAACCGCTG-----GAGGATTTAGCGCTC 1736
QY 376 IleAlaLYsAsnPheProLYsLeuSerLYsValIleSerSerLeuThrSerLYsTrpVal 395
Db 1737 GATTAAGAAACGCGACAGATGCGACGACGATGTTGGTGTGCTATGTCGCCCTATTGCC 1796
QY 396 ThrValGlnLYsGlnLYsValValVal-----AlaAlaProAlaLeuGln 409
Db 1797 ATGTACCGGCTATGTTGTGTGTCAGTTGTCGGAAAGCGCGCGCAAACTGGGT 1856
QY 410 LysGIYIleMeGIuMeGIuLeuSerGIuMeGIuGIuGIuAsnValAlaGlnPheGlnLys 429
Db 1857 AACGGCGTGAAGAAATG-----ATGGCGCAACCATTAACAGTGGTGCT 1904
QY 430 GluValGIYLYsLeuGlnAlaAlaAspMetIleSerMetPheThrGlnPheTrpGln 449
Db 1905 AACGGCTGAAA---CAGTTGGCACAAACGCGACCAAACTCTTACCAGGGGATGCA 1961
QY 450 Gln-----AlaSerLYsIleAlaSerLYsGlnThrGln 461
Db 1962 CGTATTACTAGCGCGCTGGCTATGTGGTACAGATGGCGCTCAACGAATGCTTA 2021
QY 462 SerAsnGluMet-----ThrGlnLYsAlaThrLYsGlnAlaGlnIleLeuLys 478
Db 2022 ACTAAGAAGCTGTGTAATACCTTAATAAATAGTGGCGTGGCATGGAATCAAGAT 2081

```

KAEKADNLTFRQGTANAAQONVSGEDNLSNVARLTMLAMFIEIGKNTESLQ
NDLAFENALQEGROAEMERKSAEFOETKAEETNRIMCIGKIGALLTIVYAAV
FTGASIALAAVAGVAMVADIEVKATGYSPFOOLNPIMEHVLPRIMELIKATKA
LEGIGYDKRTAMAGSIYGAITAAIAMAAYIVYAVVGAAGAAKGNLSKMGETIK
KLVPNVKQLAQNGSKLFTQGMORITSGVNGSKGLDTNLSELVNTLKNALG
MEVNTAAQAGVAGVFTKNASEALDFMLARFAMDIOOIMLKOSVEIFGEMOKVT
AELOKAMSSAVOONADASFILROSHA"
2352..3581
/gene="sipc"
/contig="sipc"
2352..3581
/gene="sipc"
/contig="sipc"
/codon_start=1
/transl_table=11
/protein_id="CAA57989.1"
/db_xref="GI:1009004"
/db_xref="SPTREMBL:Q56135"
/translaction="MLISWVGINPAAYILNHNHSEVNSSQTASQSVSAKDILNSIGISS
KYVDLGLTSLAPAPGYLTQPGTITSLKASIONTDMNODLMLANNVTKRANVY
OTLREOQAEVGFEDISGSSSAVALLAANTMLMTLNQADSKLSGKSLVSPDAK
TTASSMREGMNAISGSIOSALOGITGVGKLEKKGONERGAALKHNAKIDKLTIT
ESHSTKNVNGONSVKICAEQVDSLKSIMKKTGDATKTNIDATLKSNAQTSATSL
GTRNSKQISPEHQAILSKRLSVESDINLEONTDMTIDARKMOTDILIMKNSVT
VGGIAGASROYATQERSEQIISOVNNRVASTADSEARSSRSSTLIQEMKTMESI
NOSKASALAAIGNIRA"
3651..4670
/gene="sripd"
/contig="sripd"
3651..4670
/gene="sripd"
/codon_start=1
/transl_table=11
/protein_id="CAA57990.1"
/db_xref="GI:1009005"
/db_xref="SPTREMBL:Q56136"
/translaction="MNTIONISASPHPGIVARPOTPSASEHAEIAVNDSTEHKCTD
ITLSQAATIDQAOOTLOSTPPISEENDETERLKOQLTSSILNAKSGVLSAEON
ENRSTFSARRPYALRLMPARTISDAEIDWNSONISAGDSYLGVEVNAVYT
DEYQAFSDLSKMGWILSPKDGNTIKLVDSLSKSISSILNKYTOINKNTILFPOST
GSGMTATKAEAEOMIKELNPDCKAKASGVYVLYVDPLSKMVSIDINGISGSAI
EUDNAKYQAWQSGFKAQENLKTITQTLTKISNANSLYDNLKYLSTISSLEETAK
SFLQG"
4689..5393
/gene="sipa"
/contig="sipa"
4689..5393
/gene="sipa"
/codon_start=1
/transl_table=11
/protein_id="CAA57991.1"
/db_xref="GI:1009006"
/db_xref="SPTREMBL:Q56137"
/translaction="WTSVYRQPPYIMPGMOTETKQATNLNANLSAVRESATATLSG
ETKGOLEDFPALIKOASIDALFKCGKDEALKVEPTNENNVAGKKAIMEPAGIPESA
LNATSDSPARTLLMKVGAETAYQIITKDLKKSAPGFWLPETKRAEATLENLEKOL
DITKNNGELSKLSTNLVMOEVMPIVASCIEHNFECTLDPLTRSLQVLVKAAKA
VEALDMCHOKLQEOGT"
BASE COUNT 1584 a 1156 c 1458 g 1195 t
ORIGIN
Alignment Scores:
Pred. No.: 0.00397 Length: 5393
Score: 220.50 Matches: 116
Percent Similarity: 41.02% Conservative: 85
Best Local Similarity: 23.67% Mismatches: 182
Query Match: 9.31% Indels: 107
DB: 1 Gaps: 20
US-09-889-314-2 (1-496) x STSIPGNS (1-5393)
QY 28 GlnGlyValProGlnGlnAspLysLeuSerGlyAsnGluThrLysGlnIleGlnGlnThr 47
DB 876 CAGGCGATGATTGATGCACAAAAGAGATGGGATTCAGGTATCGAAAGATTCCAGACG 935
QY 48 ArgGlnGlyLysAsnThrGluMetGluSerAspAlaThrIleAlaGlyAlaSerGlyLys 67

```

```

DB 936 GCTGTGGGA-----GAGGCTCAGAGAGCGACGATCTCATGAAGCCACCATCAAA 986
QY 68 AspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGlyValAlaIleGly 87
DB 987 -----AGACGATVACCGCAAGAGTGTATGACGCTGGCGCC 1025
QY 88 LysGluSerSerGluSerGlnLysAla-----GlyAlaAspThrGlyVal 102
DB 1026 AAAAAATGACGACGAGCGCAAAATTAATTCGATCGGTGACCCAGCTGACCCCGCTAT 1085
QY 103 SerGlyAlaIleAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSer 122
DB 1086 GCACAACTGAAGCC-----GCC 1103
QY 123 IleGluGluAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeuSerAla 142
DB 1104 GTGAAACAGCCCGCAAAAGAACGACAGAGCGCAAGAGCCCTTAGTAAGCCACGGAT 1163
QY 143 AlAGlnMetLys-----GluValGluValValValAlaIleAlaLeuSer 157
DB 1164 GCGACGCTTAAAGCAGGACGACAGACGCCAAGCCGAAAGCGGATTAACATTTCTG 1223
QY 158 GlyLysSerSerGlySerAlaLysLeuGluThrProGluLeuProLysProGlyValThr 177
DB 1224 ACCAAATTCAGGGAACGGCTAATGCCGCTCAGATCAGATGAGTTCCAGGTTGAGCAG 1283
QY 178 ProArg-----SerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeu 196
DB 1284 GATTAATCTGCAATATGTCGCCGCCCTCACATGCTCATGTCATGTTATGGAATGTTG 1343
QY 197 GlyGluAlaThrLysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGln 216
DB 1344 GCGAAATAATRCGGAAGAACACCTGCAAAACGATTCGCTT----- 1385
QY 217 ThrAsnLysLeuGluLeuGlnLysGlnAlaIleLysIleAspLysGluArgGluGluTyr 236
DB 1386 TTCACGCCCTTGACAGGAAGGCGTCAGCGC--GAGATGGAAGAAATCGGCTGAATTC 1442
QY 237 GlnGluMetLysAlaIleGluGlnLysSerLysAspLeuGlnGlyThrMetAspThrVal 256
DB 1443 CAGCAA-----GAGACGCGCAAGCCGAGAAACGAAACGATTAAGGATGATCATC 1493
QY 257 AsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaIle 276
DB 1494 GCGAAAGTCTCGCGCGCGCTG-----CTAACCATTTCTACGCTGTGGCGCGCTGTT 1544
QY 277 PheThrCysGlyAlaGlyLeuAlaGlyLeuAlaIleGlyAlaIleValAlaValAlaAla 296
DB 1545 TTTCACGCGTGGGCGAGTGTGCGCGTGGCGTGGGACTTCGGA----- 1592
QY 297 AlaGlyClyAlaIleGlyAlaIleAlaIleAlaThrThrValAlaIleThrGlnIleThrValGln 316
DB 1593 -----ATGGTGCCCATGAATTT----- 1610
QY 317 AlaValValGlnAlaValLysGlnAlaVal-----IleThrAlaValArgGlnAlaIleThr 335
DB 1611 -----GTGAAGCGCGGACGCGGCGGTGTCGTTTATTCAGACGCGCTTAAC 1655
QY 336 AlaAlaIleLysAlaIleAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLys 355
DB 1656 CCGATTATGAGCATGNG-----CTGAAGCCGTTAATAGGAG 1691
QY 356 AlaIleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMet 375
DB 1692 CTGATTGGCAAGCGATTAACCAAGGCGTGC-----GAGGATTAAGCGCTC 1736
QY 376 IleAlaLysAsnProLysLeuSerLysValIleSerSerLeuThrSerLysThrVal 395
DB 1737 GATTAAGAAACGCGACAGATGCGACGACGATTTGTGTGCGATTTGCGCTATTGCC 1796
QY 396 ThrValGlyValGlyValValVal-----AlaAlaProAlaLeuGly 409
DB 1797 ATGCTAGCGGTATTTGTTGTTGTCGCGAGTTGTGCGGAAGAGCGGCGGCGGAATCGGGT 1856

```



```

Db 174863 AATAGCAGCAGTAACAAACAGCAACAGCAGCAACAATAGCAGCAGCAGCAAC 174804
Oy 101 yValSerGlyAlaAlaAlaAlaThrThrAlaSerAsnThr-----AlaThrIysII 117
Db 174803 AACAGTAGCAGCAGCAGCAACACACCAATTAACAAAAGCAACAGTAACAGCAATATC 174744
Oy 117 eAlaMetGlnThrSerIleGluValSerIleSerMetGluSerThrLeuGluSerIle 137
Db 174743 AGCAACAGCAACGCAATAGCAACAGCAGCAGCATCA-----GCACATATACAGCAGT 174690
Oy 137 uGlnSerLeuSerAlaAlaGlnMetLeuYsgIuValGluAlaValValAlaAlaLeuSe 157
Db 174689 AAAACAGCAACAGCAGCAGCAACACAGCAACAGCAGCAACATTAAGACA---GC 174633
Oy 157 rGlyIysSerSerGlySerAlaIleuGluSerIleuThrProGluLeuProIlyValIh 177
Db 174632 AGCAACAGCAACGTAATGACAGCAGCAGCAACACAGCAACATTAACAAAGCAACGATAGC 174573
Oy 177 rProAtgSerGluValIleGluIleGluLeuAlaLeuAlaIleGlnThrLeuG 197
Db 174572 AACGATATCA-----GCACAGCAGCAACGCAATAGCAACAAAGC 174534
Oy 197 yGluAlaThrIysSerAlaLeuSerAsnTyraIaSerThrGlnAlaGlnAlaAspGlnTh 217
Db 174533 ATCAGCAACAGCAATAGCATCAACATTAACAAACAGCAACATAGCCACAGCAGCAGCAAC 174474
Oy 217 rAsnIysLeuGlyLeuGluIysGlnAlaIleIysIleAspIysGluArgGluIuTyrgI 237
Db 174473 AGCAGCAGCAGCAACGAAATTCGACTAATTAACAGTA----- 174437
Oy 237 nGluMetIysAlaAlaGluIuGlnIysSerIysAspLeuGluGluIyThrMetAspThrValas 257
Db 174436 -----GCAGCAGCAGCAGCAATCATTCATACCATAGTAGCAACACAGTACAGCAAC 174387
Oy 257 nThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePh 277
Db 174386 ACCAGCAGCAATAGCAGCAGCAGCAATAGCA---ACAGTAATAGTAGCAACAGCAATAGC 174330
Oy 277 eThrCysGlyAlaGlyLeuAlaGlyLeuAlaIleGlyAlaIleValAlaIleAlaAlaI 297
Db 174329 AACAGTAGCAGCATCGCAACCAACAGCAGCAGCAACCAATTAAGACCCAGCAGC 174270
Oy 297 aGlyGlyAlaAlaGlyAlaAlaAlaAlaIleThrThrValAlaIleThrGlnIleThrValGlnAl 317
Db 174269 AACAGCAACAGTAGCAGCAGCAGCAACACAGTAGTAGCAGCAGCAACACAGCAAC 174210
Oy 317 aValValGlnAla-----ValIy 323
Db 174209 AATTAACAAAAGCAACAAATAGCAACAGTATCAGCAACAGCAATAGCAGCATCAGCAAC 174150
Oy 323 sGlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleIysAlaAlaValIy 343
Db 174149 AACAGCAACAAATAGCAGCAATTAACAAACAGCAGCAACAGCAACCAACAGCAGCAGC 174090
Oy 343 sSerGlyIleIysAlaPheIleIysThrLeuValIysAlaIleAlaIysAla---IleSe 362
Db 174089 AGCAACAAATAGCAGCAGCAACACAGCAACAGTAGTAGCAGCAGCAACAGCAACATTAAC 174030
Oy 362 rIlysgIy---IleSerIysValIlePheAlaIysGlyThrGlnMetIleAlaIysAsnPhEpr 381
Db 174029 AAAAGCAACAAATAGCAACAGTATCAGCAACAGCAACAGCAATTAAGAAAACAAA-----GC 173976
Oy 381 oIysIleSerIysValIleIleSerIleuThrSerIySTrIvalThrValGlyValGly-- 400
Db 173975 ATCAGCAACAGCAATAGCATCAACATTAACAAACAGCAACAAATAGCCACAGCAACAGCAAC 173916
Oy 401 -----ValValValAlaAlaProlaLeuGlyIysG 411
Db 173915 AGCAGCAGTAGCAACAAATCACTAATTAACAGTAGTAGCAGCAGCAGCAGCAGCAGCAGCAGC 173856
Oy 411 yIleMetGlnMetGlnLeuSerGluMetGlnIleAsnValAlaGlnPheGlnIlysgIuVa 431

```

```

Db 173855 AATCATTAACATAGTAGCAACAAACAGTACAGCAACACCAGCAGCAATAGTAGCAACAT 173796
Oy 431 IgIyIysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPheTrpGlnGlnAl 451
Db 173795 AGCAACAGTAATAGTAGCAACCAACAGCAGCAATAGCAACAGTAGCAGCATCGCAACAAAGAC 173736
Oy 451 aSerIysIleAlaSerIyGlnThrGlyGluSerAsnGluMetThrGlnIlysalThrIy 471
Db 173735 AACA---ATAGCAGCAGTAGTAACAAACAGCAACAGCA-----GCACCAAC 173697
Oy 471 sLeuGlyAlaGlnIleLeuIysAlaTyraAlaIleSerGlyAlaIleAlaGlyAlaH 491
Db 173696 AATAGCAGCA-----GCCGACGCAACAGCAACAGTAGTAGCAGCAGCAGC 173655
Oy 491 sIySThr 493
Db 173654 AACCACA 173648

RESULT 40
STSPGNS
LOCUS STSPGNS 5393 bp DNA linear BCT 04-OCT-1995
DEFINITION S.typhi spat, sipB, sipC, sipD and sipA genes.
ACCESSION X82670
VERSION X82670.1 GI:1009001
KEYWORDS sipA gene; sipB gene; sipC gene; sipD gene; spat gene.
SOURCE Salmonella typhi.
ORGANISM Salmonella typhi.
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
Salmonella.
REFERENCE 1 (bases 1 to 5393)
AUTHORS Hernant,D., Menard,R., Arlicau,N., Parsot,C. and Popoff,M.Y.
TITLE Functional conservation of the Salmonella and Shigella effectors of
entry into epithelial cells
JOURNAL Mol. Microbiol. 17 (4), 781-789 (1995)
MEDLINE 96111497
PUBMED 8801431
REFERENCE 2 (bases 1 to 5393)
AUTHORS Popoff,M.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-1994) M. Popoff, Institut Pasteur, Unite des
Enterobacteries, 28 rue du Docteur Roux, 75724 Paris Cedex 15,
FRANCE

FEATURES
source Location/Qualifiers
1..5393
/organism="Salmonella typhi"
/strain="Ty2"
/db_xref="taxon:601"
43..540
/gene="spat"
43..540
/gene="sipB"
/codon_start=1
/transl_table=11
/protein_id="CAA57987.1"
/db_xref="GI:1009002"
/db_xref="SPTREMBL:O57260"
/transl_table="MDYONNVSERVAEIMDAVSEGAATIKRVHGIPDDMMGTLAHA
VEFYNGRIDEETFFRFICITDFYNPDTMGLAAVCQKRFQACDIYAAFTLLK
NDYRPVFTGQCQLMRKAAROCFELVNERTEDESLRAKALVYLEALKTAETBOHS
EOEKE"
543..2324
/gene="sipB"
543..2324
/gene="sipB"
/codon_start=1
/transl_table=11
/protein_id="CAA57988.1"
/db_xref="GI:1009003"
/db_xref="SPTREMBL:O56134"
/transl_table="NYNDASSISRSRGYTONPRLAAREBGVAKNTDFLKAADKAKDY
VATKAGDLAKGTSGESAINTVGLKPPYDAAREKLSSRGQLTLGLKMTLLGLDVSLS
QLESRLAWQAMIESQKEMGIQVSKEFQALGEQADVLDEASTIKRDTASVYDAA
AKLTQQRNKLOSIDPADPGYQAQEAABAQACKFEATEAKLEALDKRATDATTVKAGTAKA

```



```

RYSHRLKTLITSTIEHOYDVIIDTPPSICMLTESALIAAOLALICATPEFYSKGLE
RLAPFIQGISRHPIILGVLTSPFNNYRKNNAATPELLOKFPCKLINTRIARDITI
SEAIHKGPFVSTAPASASEDYLTKELELFLANDI"
gene
8557..9345
/gene="TC0872"
8557..9345
/gene="TC0872"
/notes="similar to SP:P10560 GB:X07547 GB:X54450 GB:X54451
PID:581014, identified by sequence similarity; putative"
/codon_start=1
/transl_table=11
/product="viralence protein p6p6-D-related protein"
/protein_id="AAF39668.1"
/db_xref="GI:7190899"
/translation="MGNIKTLLENRFKKPTSDKMSLAKKRLEGELSPFLNFTPKL
SSQERAKRROLLEYSFSPNEISKHDLOQJLHSAQVKQJHDAVLLHGERIKKVELL
TSYRGASAMALLTYGNKOPYNFLVLYELSTPDSLKLEKLPRQAVYTLARE
GTQEKKEA1IRNYOGETKGELLEIRKEFPILPTDRRHTPLSOKAFETTLAKGKTKLRO
CTDVSHEELIALLEKLIKRLQKVTNLSNTKVS RMDNEI0NSRN"
9314..9865
/gene="TC0873"
9314..9865
/gene="TC0873"
/notes="conserved hypothetical protein; identified by
glimer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF39669.1"
/db_xref="GI:7190900"
/translation="MTTKSKTLEIDNNFLLLEGNLKRIFATPIGTTTREFRONYFN
CAQGOELANFLFEMLINGKILLOELPAGOKOASOILVOFMMILIRVADIIHGERFIN
FTSDMLAOEBCVFLNLSRYDGOEPLMTVONTCHILIRLLSLRLEAKNP1GKEK
NLOEVEDIDLSRAHFEELTKSM"
BASE COUNT      3291 a      2262 c      1904 g      2526 t
ORIGIN

Alignment Scores:
Pred. No.:      4,84e-57      Length:      9983
Score:          1223.50      Matches:      263
Percent Similarity: 69.98%      Conservative: 89
Best Local Similarity: 52.29%      Mismatches: 122
Query Match:      51.65%      Indels:      29
DB:              1          Gaps:      9

US-09-889-314-2 (1-496) x AB002353 (1-9983)
QY      4 MetSerIleSerSerSerSerGlyProAsp-----AsnGlnIlysAsnIleMetSerGln 21
      1316 ATGTCTCTTCAATCTCTCTGCTCTGTGACAGACACATCTAAGAGTGTATCTCAA 1375
QY      22 ValLeuThrSerThrProGlnGlyValProGlnIAspIysLeuSerGlyAsnGlnThr 41
      1376 GTCATTACATCGACTCTCAAGCGCTTCTTAATGACATTAAGCTGACACATTCAGGTT 1435
QY      42 LysGlnIleGlnGlnThrArgGlnGlyLys---AsnThrGlnMetGlnIysSerAspAlaThr 60
      1436 AAGACAGGTCCACAGACTGACAGAACCGAGACGACCTGACATGGAAGAAACGATTACGCT 1495
QY      61 IleAlaGlyAlaSerGlyLysAspIysThrSerSerThrThrLys-----ThrGln 77
      1496 ATTCGCTGGAGCGGCTGGGAAGATCGTCATCTCTGTTCCCAATGGAAGCCCAAG 1555
QY      78 ThrAlaProGlnGlnGlyAlaAlaGlyLysGlnSerSerGlnIysAlaGly 97
      1556 GTCATTGAGACAGCAAGATTAGCCCTGGGAAGAAACGCC-----1597
DB      1556 GTCATTGAGACAGCAAGATTAGCCCTGGGAAGAAACGCC-----1597
QY      98 AlaAspThrGlyAlaSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrIle 117
      1598 -----TCTGCAGACGACAGCATTCCTTAACAGAACCTTCCACAGAGCA 1642
DB      1598 -----TCTGCAGACGACAGCATTCCTTAACAGAACCTTCCACAGAGCA 1642
QY      118 AlaMetGlnThrSerIleGlnGlnAlaSerIysSerMetGln---SerThrLeuGlnSer 136
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111

```

```

DB      1643 GCTTCTCAAAAACGCTAGACAGATACCAACAATCTTGAACATCTCTTAAGTCC 1702
QY      137 LeuGlnSerLeuSerIleAlaGlnMetIysGlnValGlnAlaValAlaAlaLeu 156
      111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
DB      1703 CTCTCCCTCCGTAGATCGCCCAACACAAAGAAATTCAGAACATGTGCTTCCGCAAG 1762
QY      157 SerGlyLysSerSerGlySerAlaLys---LeuGlnThrProGlnLeuProIysProGly 175
      1763 AGATCTACTACAGACATTCCTTAAGAACTAGACTCTCCGAGCTACCCAGCTTACA 1822
DB      1763 AGATCTACTACAGACATTCCTTAAGAACTAGACTCTCCGAGCTACCCAGCTTACA 1822
QY      176 ValThrProArgSerGlnValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThr 195
      1823 ACAGACGCTCCGCAAGATGATGTAATGCAATCAGCTTATGCGAAGCAATTCAGCC 1882
DB      1823 ACAGACGCTCCGCAAGATGATGTAATGCAATCAGCTTATGCGAAGCAATTCAGCC 1882
QY      196 LeuGlyGlnAlaThrIysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAsp 215
      1883 CTTCGGAGAGCTTCAGCAGACAGCTTGGAAATTCCTCAAGACGCAAACTCAAGCTCC 1942
DB      1883 CTTCGGAGAGCTTCAGCAGACAGCTTGGAAATTCCTCAAGACGCAAACTCAAGCTCC 1942
QY      216 GlnThrAsnLysLeuGlyLeuGlnIlysGlnAlaIleLysIleAspIysGlnArgGlnGlu 235
      1943 AATATGACAAATATGCTCTGGAATCTCAAGACTCAAAATTCACAAAGACAGCAAGAG 2002
DB      1943 AATATGACAAATATGCTCTGGAATCTCAAGACTCAAAATTCACAAAGACAGCAAGAG 2002
QY      236 TyrGlnGlnMetLysAlaAlaGlnGlnIysSerIysAspLeuGlnIlyThrMetAspThr 255
      2003 TTCCAAAAATTCGACAGATCCGAAAAAAGCTGGCAAC---AATCCACAGATGATACC 2059
DB      2003 TTCCAAAAATTCGACAGATCCGAAAAAAGCTGGCAAC---AATCCACAGATGATACC 2059
QY      256 ValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAla 275
      2060 GTTAAACAAGTGATGATGTGAGTGGAGTCTGCTATCATCTGATTTCTGTTGTAGCAGCA 2119
DB      2060 GTTAAACAAGTGATGATGTGAGTGGAGTCTGCTATCATCTGATTTCTGTTGTAGCAGCA 2119
QY      276 IlePheThrCysGlyIysAlaGlyLeuAlaGlyLeuAlaAlaGlyLysAlaValGlyAlaAla 295
      2120 CTAATTACTTGGCGCTGCTGTTTAATCGGACAGCAGCAGCAGGAGCTTACA---GCAGCA 2176
DB      2120 CTAATTACTTGGCGCTGCTGTTTAATCGGACAGCAGCAGCAGGAGCTTACA---GCAGCA 2176
QY      296 AlaAlaGlyIysAlaAlaGlyAlaAlaAla-----AlaThr 307
      2177 GCAGCGGAGGCCACAGAGGCTTACAGCAGCAGCAACGACAGCACTTGTGGCAACA 2236
DB      2177 GCAGCGGAGGCCACAGAGGCTTACAGCAGCAGCAACGACAGCACTTGTGGCAACA 2236
QY      308 ThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValIysGlnAlaValIle 327
      2237 ACAGTTCGCGCCCAAGATTACGATGACAGCAGCATCAGCGGTTAAGCAAGTATCATTT 2296
DB      2237 ACAGTTCGCGCCCAAGATTACGATGACAGCAGCATCAGCGGTTAAGCAAGTATCATTT 2296
QY      328 ThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaAlaValIysSerGlyIleLys 347
      2297 CAAGGGGTCAAGAGGCTATTGTCCAAAGCATTAAGCAGGACTTTAAAAAGGATCATG 2356
DB      2297 CAAGGGGTCAAGAGGCTATTGTCCAAAGCATTAAGCAGGACTTTAAAAAGGATCATG 2356
QY      348 AlaPheIleLysThrLeuValIysAlaAlaIleAlaLysAlaIleSerIysGlyIleSerLys 367
      2357 CAAGCATTTAAACAGGCTATTAAGAGCTGCTTAAGAACTTACCAAAAAATATTGGTAA 2416
DB      2357 CAAGCATTTAAACAGGCTATTAAGAGCTGCTTAAGAACTTACCAAAAAATATTGGTAA 2416
QY      368 ValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProIysLeuSerIysValIle 387
      2417 ATTTTCAACACAGAGAAATAATGCCGTAGTAATTCCTTCCCTTAACCTCAACAAAGTAG 2476
DB      2417 ATTTTCAACACAGAGAAATAATGCCGTAGTAATTCCTTCCCTTAACCTCAACAAAGTAG 2476
QY      388 SerSerLeuThrSerIysThrPvalThrValIysValGlyValAlaValAlaAlaProIle 407
      2477 AACACTCTGGAGATTAATGAGTAACTTGAAGATTGAGACCTGACAGAGTTCCCAA 2536
DB      2477 AACACTCTGGAGATTAATGAGTAACTTGAAGATTGAGACCTGACAGAGTTCCCAA 2536
QY      408 LeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnIysAsnValAlaGlnPhe 427
      2537 CTTCGTAACAGAGTTTGGAAATATCGACGCTATTCGACATGCAAAAAATCTGTGAAGATT 2596
DB      2537 CTTCGTAACAGAGTTTGGAAATATCGACGCTATTCGACATGCAAAAAATCTGTGAAGATT 2596
QY      428 GlnIysGlnValGlyIysIysGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPhe 447
      2597 CAAAAAGAGTCCGAGACTCTTACAGCCCAATCGAATGTGAAGCGCTTACACTTGTTC 2656
DB      2597 CAAAAAGAGTCCGAGACTCTTACAGCCCAATCGAATGTGAAGCGCTTACACTTGTTC 2656
QY      448 TrpGlnGlnAlaSerIysIleAlaSerIysGlnThrGlyLysSerAsnGlnMetThrGln 467
      2657 TGGCAACAGCAAGTAAATGCTGCAAAACCAAGCAAGCCCTGAGCAACACAAACAG 2716
DB      2657 TGGCAACAGCAAGTAAATGCTGCAAAACCAAGCAAGCCCTGAGCAACACAAACAG 2716
QY      468 LysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaThrAlaAlaIleSerGlyAlaIle 487
      2717 CAAGCGGCCAAAAACAGAGACTCAATATAGCAAAAGCTTGTCCGCAATTAAGTGGCGCTCA 2776
DB      2717 CAAGCGGCCAAAAACAGAGACTCAATATAGCAAAAGCTTGTCCGCAATTAAGTGGCGCTCA 2776

```

TITLE	Bowman,C., Dodson,R., Gwin,M., Nelson,W., DeBoy,R., Kolonay,J., McClarty,G., Salzberg,S.L., Eisen,J. and Fraser,C.M. Genome Sequences of Chlamydia trachomatis Mohn and Chlamydia pneumoniae AR39
JOURNAL	Nucleic Acids Res. 28 (6), 1397-1406 (2000)
MEDLINE	20150255
PUBMED	10684935
REFERENCE	2 (bases 1 to 9983) Read,T.D., Bruham,R., Shen,C., Gill,S.R., Heidelberg,J.F., White,O., Hickey,E.K., Peterson,J., Umayam,L.A., Uterback,T., Berry,K., Bass,S., Linher,K., Weidman,J., Khouri,H., Craven,B., Bowman,C., Dodson,R., Gwin,M., Nelson,W., DeBoy,R., Kolonay,J., McClarty,G., Salzberg,S.L., Eisen,J. and Fraser,C.M.
TITLE	Direct Submission Submitted (01-MAR-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA
JOURNAL	On Jun 1, 2000 this sequence version replaced gi:7190891.
COMMENT	Location/Qualifiers
FEATURES	1..9983 organism="Chlamydia muridarum" db_xref="taxon:83560" note="synonym: Chlamydia trachomatis Mohn" 152..892 gene="TC0865" 152..892 gene="TC0865" 152..892 gene="TC0865" /note="similar to GB:M26405 SP:P21207 PID:155451; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="type III secretion chaperone SycD" /protein_id="AF39661.1" /db_xref="GI:7190892" /translation="MLVEVIEFELGIIAMSTPSSNNKRPSPASFNKKSRSLATIAA OKAKADLEPKYVPPEBEETKVLEIIGGLQOILGLDVLEEIVTAAV FYSGCKYREAIIFOLITASKPCQCYKIIGSCCHQIKMDEAPGFLEADNPEN PIPPYITADSLMDLPESONFLDTITMCNKNRPTVKLKERCSIMKOSLDAMLKE KRASASSSEKSTTTAKRSSPNKKRAK"
CDS	920..1279 /gene="TC0866" 920..1279 gene="TC0866" /note="conserved hypothetical protein; identified by Glimmer2; putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAF39662.1" /db_xref="GI:7190893" /translation="MSKKHKPNKRRSSSKSQYKPAHLSTKEAPALQDIOTIFSOD LPLAQMSEMOMODEKOLAKMMALSMGLDSLPIETLTGKLINNKBEAQLSOEISSIFL GLRLITFAVNKHIADEK"
gene	1316..2791 /gene="TC0867" 1316..2791 gene="TC0867" /note="conserved hypothetical protein; identified by Glimmer2; putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAF39663.1" /db_xref="GI:7190894" /translation="MSLSSSSSSDSNLRSVLSTPOGTTPPGVPMADKLTIDNQAVO OTRONPDLIMESDALACAGACKDRASSVSQMEGEVTBOCGLAGKEPASAISTRTST ONASTGAASOKRVEDTNKSLESSLSSVDASQLOEIONTVASATRNSTNSIKNL DPTEPLPPTTPRODWELISLAKAITALGSTQALENFSTQTOQASNMRKMLE OGKIDEREFOQLQEOIKKAGNNSTMPTVNMVGIVALTVAITSVAALFTCGIL IGTAAGAATAAGATAGATAATTATSVATTAGOVTMQAOWAKOAIIIOGVKAI VOAIKQIKKGIMGAIKOATKAATKTLTICKNTICKIPFGTKGNVSAKSPNISKMTNGLS KWTLTGALITANPOLVSGIGNIQLSDMOKDLAKIOKEGALTAAOSEMKKATLTWQQ ASKIAAQITSPSETQQAKTGAQTAKALSATSIGLAAAA"
CDS	2822..4132 /gene="TC0868" 2822..4132 gene="TC0868" /note="conserved hypothetical protein; identified by Glimmer2; putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAF39664.1" /db_xref="GI:7190895" /translation="MTTYGRCNFPDPSLLADLTQNANSASDPKNCMAAKTQOOEEBA SFEDLIODSGCKSKASTSOTSRSKSKRKSGGGTTTTVAOAOTPAOVAOVN ARSGFNNGSGASITSPSPNPYNGVIRKMNQTLITLAASAKSSRPOO QNOATINOYAMAPEITNAIRTOAHNOAAATELOAKOSITGITTAGVSYGGCIIS ASKSLDGKLSAAFEITNAITRAASSASAAKATVNALDIVANVAATAGKASGAASAA SSAATLTIONMEDTASKTLSTQTSKATGTFENALNPNMSEKISRGINLVKTIQGPAA AOPAGRALSSAMSISOMVGLTAGVDGTTGGTIGAOVNEORAGAMEARAEEKLSLN SVOSYASDAQOQOEOSOFNSALQTLQSIADSSLQTTGAMFN"
gene	complement(4206..5162) /gene="TC0869" complement(4206..5162) gene="TC0869" /note="conserved hypothetical protein; identified by Glimmer2; putative" /codon_start=1 /transl_table=11 /product="conserved hypothetical protein" /protein_id="AAF39665.1" /db_xref="GI:7190896" /translation="MWLYRFAPNAVYEPARFLPLNMFIVSSSPATSKSAMDAAPL FVTSGRNVLAGIYLREGLLFKRESFRPLROAYMPVLVLSYGFIPLYNLETTIGQGIS SSKACEFTYGSFPTAFCSTVOLREVYVMKUGLSLQVSTLVYLLFNGGEDVSEWGK WOLGMPELLIVATCLSSYGWTLLRLEROCESLITAINVANVIAGTILASHSATIT EWNPJPVENPVLFLOSIGALVFSMLICYNFLFKILSFESTFISFCNIWMPFAFSE FGMLLGESPFPPLLFAVGFVWICRLIYHEBERQGYVLSBP"
CDS	5409..7316 /gene="TC0870" 5409..7316 gene="TC0870" /note="similar to GB:M36593 SP:P18256 PID:143764 PID:1565235 GB:A009126; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="threonyl-tRNA synthetase" /protein_id="AAF39666.1" /db_xref="GI:7190897" /translation="MTHVCNQAEFLPGASANDLASRMKNSHCFAGALINDOEKDL STLKDDGYLFLTWDPOGREFLHTSHNTIAQAVLRKMPSAQFTICPYIDQGPFYTD FANLISISEDFEPAITEMARALAEKLIASROYFSKEETFLAFSNSPKIELITELPEL EATISAYKGEFFPALRGPHLPSTAPVKAFKLRTSAWRQDSKEPILRIYIAEPFEL TTKELEHLHOLEEAKRDHRVLTGLDLFSQSSACMPFHPRGVAINNALIDYWKM RLHOLAGYOQIOTPOLKSRRELMEISGHMENYKENYTTLVEDEVAIKPMKPCGMXYL XKQLUSIREPFLRIAEIGHVRHELGSALISMVTRHODDAHVELTLPPOVEERTL NITNLVSELCTGGLDYHLELTRPQGTGSDDIMELATEALKALKYSKOPTIIS GEGAFYGPKLDIHVKAIINKTWCCGTIQDMFLPERFDLKTTPNGEKSTPIMLRAL FGSIERFLGLILEHFGRPLMISPENHVIITVAARHRAQELAKEYTPOOGIATYADV QASESVSKIRIANQNMVNMITIGDKELETO LAVFRDNRLNDISVEHFTWLTILE ELRSLSTSLSL"
gene	7785..8552 /gene="TC0871" 7785..8552 gene="TC0871" /note="similar to GB:D26185 SP:P37522 PID:467381 PID:580996 GB:A009126; identified by sequence similarity; putative" /codon_start=1 /transl_table=11 /product="Para family protein" /protein_id="AAF39667.1" /db_xref="GI:7190898" /translation="MKTIVNSFKGTAKTSTTLHLGALAQYHKARVLLIDPDQAQAN LTFAGLIDPCQYDLSLAVVILQGEKQISSEVIRSIDDSSGLDIRADPWLTLEVENVSGSLAAD

CDs 8492..10399
/gene="chrs"
/codon_start=1
/transl_table=11
/product="threonyl tRNA Synthetase"
/protein_id="AAC68183.2"
/db_xref="GI:6578108"
/translation="MIHYTCNOAELEPEGASAMDLANKMKSHCFAGALINDKERL
STTLQDGVLEFLTWDDPKGREIFLHTSAHILAOAVLRMPASOPTIGVIDQGFYD
FANLISSEDPALIEAMAKTIAEEKPILSRQVPEPKALAEFSONPKRAELIAEPE
EVEISATYQGEPLDLCRPHLPSTAPVAFKILRTSSAVKWDPSRESILRTYGSFP
TTKLEKHLHOLEEAKRRDHRVLTGKLDLFSQOTCSAGPEFHPRGMVWVALVDYWK
RLHORAGYQOQIOTPOLMNRLEMEISGHENYKENVYTLVDEEDYALRPMCGCLY
YKTOHSYREPLRAIEIGHVHRHELSCALSGIMLVRTFHODAHVFLPEVEETL
NILNVLSELYGFEGLYELHLESTREPGCTIGSDMLMELATKALRYLKSOKPFIIS
GEGAFYGPKIDIHVKDAINRMQCTTOLDMLPRPLKRYNACGEKSTIMLRAL
FGSIRERGLIEHFKGRPLMLSPENRITTVADRHEARAOELKHFSONGITVSD
SSNESVSKIRINQAMQVNYMTIGDKLETHLLAVRTRDNKVLMDIAVEQSHVILE
ELRSLTNPST"

BASE COUNT 3381 a 2527 c 2095 g 2650 t
ORIGIN

Alignment Scores:
Pred. No.: 1.51e-57 Length: 10653
Score: 1233.50 Matches: 268
Percent Similarity: 70.14% Conservatve: 82
Best Local Similarity: 53.71% Mismatches: 124
Query Match: 52.07% Indels: 25
DB: 1 Gaps: 10

US-09-889-314-2 (1-496) x AE001328 (1-10653)

OY 4 MetSerIleSerSerSerGlyProAsp-----AsnGlnLysAsnIleMetSerGln 21
DB 4313 AGTCCCTTCATCTTCTTCTGCTCTCCGATAGCAGCACTTAGAGATGCTTGCGCA 4372
OY 22 ValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThr 41
DB 4373 GTCATAGCTTCCGACCTCCGACGCTTCTTAATGACAGATAAATTAAACGACATCAGGTT 4432
OY 42 LysGlnIleGlnGlnThrArgGlnGlyLys---AsnThrGluMetGluSerAspAlaThr 60
DB 4433 AAGCAGATTCAACAGACGACGACAAATCGCATGACCTAAGCTGGAAGCGATGCGCT 4492
OY 61 IleAlaMetGlnSerGlyLysAspLysThrSerSerThrThryls-----ThrGlu 77
DB 4493 GTTCCGCGAGCTGCTGGAAAGATCGCGACGCTTCTTCAATAATAGAGACAGCAAGA 4552
OY 78 ThrAlaProGlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGly 97
DB 4553 CTTATAGACGACAGAGATAGCTGCGAGGAAGAAACT-----GCATCT 4597
OY 98 AlaAspThrGlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLys--- 116
DB 4598 GCGGAG-----GCGACATCCCTAACCAGCAAGCCGACATTAAGA 4636
OY 117 IleAlaMetGlnThrSerIleGlnGlnAlaSerLysSerMetGlu---SerThrLeuGlu 135
DB 4637 GCTAGCTCGCAACATGCTATAGAGATAGCTAGCAAAATCTTTAGAGCTATCTTTTAAGT 4696
OY 136 SerLeuGlnSerLeuSerAlaAlaGlnMetLysGluValGlnAlaValAlaAlaAla 155
DB 4697 TCGTTGTCTATCTGTAGATCCACGACATCTACAAAGAAATTCAGATCGATCTGTGCT 4756
OY 156 LeuSerGlyLysSerSerLysSer---AlaLysLeuGlnThrProGluLeuProLysPro 174
DB 4757 ATGGGTGCTACTACGACGCTTCTTGACGAACTTAGAAACTCCAGACTACCAACCT 4816
OY 175 GlyValThrProArgSerGlyValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGln 194
DB 4817 TCAAGACACCTCGTCAAAAGTAATGAAATTTAGCTTGACATTAACAAAGCAATTTACC 4876
OY 195 ThrLeuGlyGlnAlaThrLysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAla 214

DB 4877 GCTCTGAGAGTCACGACGACGACGATTCGGAACCTTCACAAAGTACGACATCT 4936
OY 215 AspGlnThrAsnLysLeuGlyLysGlnAlaIleLysIleAspLysGluArgGlu 234
DB 4937 GCGAATGACAAACAAATGCTCTAGAAATCTCAAGGCTCTAAATGATTAAGGCTGAA 4996
OY 235 GluTyrGlnGluMetLysAlaAlaGlnGlnLysSerLysAspLeuGlnLysThrMetAsp 254
DB 4997 GAGTTCAAAAAATGCAAGAGATCCGCAAAATCT---GGAACCACTCTACCATGCAAT 5053
OY 255 ThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAla 274
DB 5054 ACCGTTAAACAAAGATGATGATTTGGGCTTACCGCGGTATTAAGTGTATCTGTAGATATCC 5113
OY 275 AlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaValAlaGlyAla 294
DB 5114 GCATTTTCACTTGGCTGCTTGGCTGATTCGGAACGCTGTCAGAGCCACAGCAGCC 5173
OY 295 AlaAlaAlaGlyValAlaAlaGlyAlaAlaAla-----AlaThrThrValAlaThr 311
DB 5174 GCGGCTGAGCTACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 5233
OY 312 GlnIleThrValGlnAlaValAlaValAlaValAlaValAlaValAlaValAlaValArg 331
DB 5234 CAAGTACATATGCAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 5293
OY 332 GlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyLysAlaPheIleLys 351
DB 5294 CAGGCTATGCTCCAACTATTTAAACAAAGGATTTAAACAAAGGATTTAAACAAAGGATTTAA 5353
OY 352 ThrLeuValLysAlaIleAlaLysAlaIleSerLysGlyLysSerLysValPheAlaLys 371
DB 5354 CAAGCTGTTAAGGCGGCTGTGAACACCTTCTCTAAACGTTGGTAAATTTTTCACGCA 5413
OY 372 GlyThrGlnMetIleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThr 391
DB 5414 GGGAAATATGCTGTATGCAATATGCTTCCCTAAACCTCCAAAGTTATCAACATTTGGGA 5473
OY 392 SerLysThrValThrValAlaGlyValAlaValAlaValAlaAlaProAlaLeuGlyLysGly 411
DB 5474 AGTAAATGGGTAACTTACAGATGAGAGCTCTTACGACGCTTCCAACTGATTCGCGG 5533
OY 412 IleMetGlnMetGlnLeuSerGlnMetGlnGlnAsnValAlaGlnPheGlnLysGluVal 431
DB 5534 ATTACTAGTCTGAGCTGTGACATGACATGCAAGAAAGAACTGCCCCAAATTCAAAAGAAAGTC 5593
OY 432 GlyLysLeuGlnAlaAlaAlaAlaAspMetIleSerMetPheThrGlnPheProGlnAla 451
DB 5594 GGAAGCTCTCAGACGCTCAATCTGAAATGATGAAGAAAGCTTTCACATTTTTCGGAACAAGCA 5653
OY 452 SerLysIleAlaSerLysGlnThrGlyLysGluSerAsnGlnMetThrGlnLysAlaThrLys 471
DB 5654 AGTAAATATGCTGTATGCAATATGCTTCCCTAAACCTCCAAAGTTATCAACATTTGGGA 5713
OY 472 LeuGlyAlaGlnIleLeuLysAlaThrAlaAlaIleSerGlyValAlaIleAlaGlyAla 490
DB 5714 ACCGAGCTCAGATACGAAAGCTTTTCCGCAATTAATGAGGCGCTTACCGCGCGCA 5770

RESULT 38
AE002353 9983 bp DNA linear BCT 26-MAY-2000
LOCUS Chlamydia muridarum, section 80 of 85 of the complete genome.
DEFINITION AE002353 AE002160
ACCESSION AE002353.2 GI:8163337
VERSION
KEYWORDS
SOURCE
ORGANISM Chlamydia muridarum.
Chlamydia muridarum.
Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
REFERENCE 1 (bases 1 to 9983) Shen, C., Gill, S.R., Heidelberg, J.F.,
White, O., Hickey, E.K., Peterson, J., Umayam, L.A., Utteback, T.,
Berry, K., Bass, S., Linher, K., Weidman, J., Kouri, H., Craven, B.,

OY		144	GlnmetLysGluValGluAlaValValValAAlaIalaLeuSerCylLysSerSerCylSer	163
Dd		904	CAAATGAAGAAGTGTCAAGCGGGTTGTCTGCCCTCCACAGAAAAATTTCGGGGTTCC	963
OY		164	AAlaLysLeuGIUThPrOgLUleuProLYSProGLyVALThPProArgSerCIuVaiIle	183
Dd		964	GCAAATTGGAAAACAACCCTGCAGCTCCCACCAAGCCCCGGGGTGACACCCAAGATCACAGGTTATC	1023
OY		184	GluilegllyLeuaIaleLuAlaLySaIAIllegInTrhleugllyGlualatHrLysSeRaI	203
Dd		1024	GAATTCGACTCGCGCTTCCTTAAGCAATCATGACATTGGGAGAACCAATAAATCTGCC	1083
OY		204	LeuSerAsnTYryALISerThrGlnAlaaglinAlaaRgrInThAsnLysLeuGluGlu	223
Dd		1084	TTCATTAACTATGACAATPCACAMACACACAGACCAAAACAATAAATACGTAGTAGNA	1143
OY		224	LysGINalIalleLysIIeasPLysGUarAgUGJuTytyrgINgleIMetLyslaAIagu	243
Dd		1144	AAGCAAGCATTAATAATCGATTAATAAGACAGAGAATACCAAGAGATGAAAGCTCCGHA	1203
OY		244	GlnLysSerLysASpDeuGluGluTYthrMetasprrThvaLnTHvalMetlleaIa	262
Dd		1204	CAGAAGTCTAAAGATCTCGAAGACAATGATGTACTGTCAATACGTGTATGATCGCG	1260
RESULT		36		
LOCUS	E12539		1296 bp DNA linear PAT 27-APR-1998	
DEFINITION			DNA encoding a fusion protein of DHFR and antigen peptide of Chlamydia pneumoniae.	
ACCESSION	E12539			
VERSION	E12539.1 GI:3251372			
KEYWORDS	JF 1997015243-A/2.			
SOURCE	unidentified.			
ORGANISM	unclassified.			
REFERENCE	1 (bases 1 to 1296)			
AUTHORS	Izutsu H., Obara K. and Matsumoto A.			
TITLE	MESURING METHOD FOR ANTICHLAMYDIA PNEUMONIA ANTIBODY AND REAGENT THEREFOR AS WELL AS DIAGNOSTIC MEDICINE FOR CHLAMYDIA PNEUMONIA INFECTION Patent: JP 1997015243-A 2 17-JAN-1997; HITACHI CHEM CO LTD			
JOURNAL COMMENT				
OS	None			
OC	Artificial sequences.			
PN	JP 1997015243-A/2			
PD	17-JAN-1997			
Pf	15-MAR-1996 JP 1996058609			
PR	28-APR-1995 JP 95P 106009			
PI	IZUTSU HIROSHI, OBARA KAUKHIKO, MATSUMOTO AKIRA PC G01N33/56B,C12N1/21,C12N15/02,C12N15/09,C12P21/02,C12P21/08, PC G01N33/53,			
PC	PC G01N33/53,I,G01N33/57I//A61K49/00,C07K14/245,C07K14/39, PC C07K16/12,C07K19/00.			
PC	(C12N1/21,C12R1:19),(C12N15/09,C12N1:01),(C12P21/02,C12R1:19);:			
CC	Strandedness: Double;			
CC	Topology: Linear;			
FH	Key Location/Qualifiers			
FT	source 1..1296			
FT	malf-peptide 1..1296			
FT	/product='fusion protein of DHFR and antigen peptide'			
FT	misc-feature 1..480			
FT	/note='DHFR coding region'			
FT	misc-feature 484..1296			
FT	/note='antigen peptide coding region'.			
FEATURES source	Location/Qualifiers			
	1..1296 /organism='unidentified' /db_xref='taxon:32644'			
BASSE COUNT ORIGIN	407 a 299 c 325 g 265 t			

Alignment Scores:			
Pred. No.:	2,866-59	Length:	1296
Score:	1245.00	Matches:	259
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	52.55%	Indels:	0
DB:	6	Gaps:	0
US-09-889-314-2 (1-496) x E12539 (1-1296)			
QY	4	MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu	23
DB	484	ATGCTATTTCATCTCTTCTCAGACNCTACATCAATCAAAAAATATCATCTCTCAAGTTCTG	543
QY	24	ThrSerThrProGlnGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln	43
DB	544	ACATCGACACCCCCAGGGGGTGGCCCCCAACAAATATGAGCTGTGGCAACGAACCAAGCAA	603
QY	44	IleGlnGlnThrArgGlnGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly	63
DB	604	ATACGACCAACCGCTCAGAGTAAACACTGAGATGGAAGAACGATGCCATATGGCTGGT	663
QY	64	AlaSerGlyLysAspLysFthrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly	83
DB	664	GCTTCTGGAAAAAGACAACAACTTCTCCGACATCAAAAAACAGAAAGCTCCACAAACAGGA	723
QY	84	ValAlaIleAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThrGlyValSer	103
DB	724	GTTCGTGCTGGGAAAGATCCTCAGAAAGTCAAAAGCGAGGTGTGAATCGGAGTATCA	783
QY	104	GlyAlaAlaIleAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle	123
DB	784	GGAGCGGCTGCTACTACAGCATCAAAATCTGCACAAAAATTTGCTATGCACACTCTATT	843
QY	124	GlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla	143
DB	844	GAAGAGCGCAGACAAAGATGAGAGCTCACTTTAGAGTCAATTCAAAGCCTCAGTCCGCG	903
QY	144	GlnMetLysGlnValGlnAlaValAlaValAlaIleLeuSerGlyLysSerSerGlySer	163
DB	904	CAAAAGAAAGAAAGTCGAAGCGGTTGTTGTTCTGCTCCCTCAGGGAAAAAGTTCCGGTTTC	963
QY	164	AlaLysLeuGlnThrProGlnLeuProLysProGlyValAThrProArgSerGlnValIle	183
DB	964	GCAAAATTTGGAACACCTCGAGCTCCCAAGCCGGGGTGAACCAAGATTCAGAGTTATC	1023
QY	184	GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlyGlnAlaThrLysSerAla	203
DB	1024	GAATTCGGACCTCGCCTTGTCAAGCAATTCAGACATTCGAGAGCAAGCCACCAAAATCTGCC	1083
QY	204	LeuSerAsnThrAlaSerThrGlnAlaIleAlaLysGlnThrAsnLysLeuGlnLeuGln	223
DB	1084	TTATTTAACTATGCAAGTACCAACACACACAGCCCAACAAATAATTAACATVAGGCTTGAA	1143
QY	224	LysGlnAlaIleLysIleAspLysGlnLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	243
DB	1144	AAGCAAGGATTAATAATTCGATTAAGAACAGACAGATATCCAAAGATGAAGCTGCCGAA	1203
QY	244	GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAla	262
DB	1204	CAGAGTCTAAAGATCTCGAAGGAACATGATCTGATCATCTGATGATGATGCGC	1260
RESULT 37			
LOCUS	AE001328	10653 bp	DNA linear BCT 30-Oct-2000
DEFINITION	Chlamydia trachomatis section 55 of 87 of the complete genome.		
ACCESSION	AE001328	AE0013273	
VERSION	AE001328.1	GI:3329015	
KEYWORDS			
SOURCE	Chlamydia trachomatis.		
ORGANISM	Chlamydia trachomatis.		
	Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.		

US-09-889-314-2 (1-496) x AR122115 (1-1296)

QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23
 DB 484 ATGCTATTTCATCTCTTCTGAGCCTGACAAATCAAAAAATATCATGCTCAAGTTCTGT 543

QY 24 ThrSerThrProGlnGlnValProGlnGlnAspLysLeuSerGlyAsnGlnIleThrLysGln 43
 DB 544 ACATCGACACCCAGGCGGTGCCCCAACAGATAGCTGTGTGCGAACGAAACGACAA 603

QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63
 DB 604 ATACAGCAACACGCTGAGGTGTAACAAACAGTGAAGGCGATGCCACATATGCGGT 663

QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83
 DB 664 GCTTCTGAAAAGACAAATCTCTCGACTACAAAAACAGAACGCTCCACAAACAGGGA 723

QY 84 ValAlaIleAlaGlyLysGlnSerSerSerGlnLysAlaGlyAlaAspThrGlnValSer 103
 DB 724 GTTCTGCTGGGAAAGATCTCTCAAGATCAAAAGCGAGTGTGATCTGAGATATCA 783

QY 104 G1ValAlaIleAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
 DB 784 GAGCGGCTGCTACTACAGCATCAATACATGCAACAAAANTGCTATGCAACCTCTATT 843

QY 124 GlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143
 DB 844 GAAGAGCGGACCAAAAGTATGAGTACCTTACAGTCACTTCAAAACCTCAGTGCAGCG 903

QY 144 GlnMetLysGlnValGlnAlaValAlaValAlaLeuSerGlyLysSerSerGlySer 163
 DB 904 CAATGTAAGAGAGCGAAGCGGTGTGCTGCTCTCAAGGAAAGTTCGGGTTC 963

QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnValIle 183
 DB 964 GCAAAATTTGGAAACACTGAGCTCCCAAGCGGGTGACACCAAGATCAGAGATTATC 1023

QY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThrLysSerAla 203
 DB 1024 GAATTCGAGCTCGCGCTTGTCTTAAAGCAATTCAGACATTTGGAGAGGACACAAATCTGCC 1083

QY 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnIleGln 223
 DB 1084 TTATCTACTATGCAAGTACACAGCAACAGCAACCAAAATATACATGATCTGAAA 1143

QY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnIleThrGlnGlnMetLysAlaAlaGln 243
 DB 1144 AAGCAACGATTAATAATCGATTAAGAACGAGAACATACCAAGATGAAAGCTGCGGAA 1203

QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAla 262
 DB 1204 CAGAAGTCTAAGATCTCGAAGAACATGATGATCTGATGATGATGCGC 1260

RESULT 35
 LOCUS E12526 1296 bp DNA linear PAT 27-APR-1998
 DEFINITION DNA encoding a fusion protein of DHFR and antigen peptide of Chlamydia pneumoniae.
 ACCESSION E12526
 VERSION E12526.1 GI:3251359
 KEYWORDS JP 1997009976-A/2.
 SOURCE unidentified.
 ORGANISM unidentified.
 REFERENCE 1 (bases 1 to 1296)
 AUTHORS Iizutsu H., Obara K. and Matsumoto A.
 TITLE DIHYDROFOLIC ACID REDUCTASE-CHLAMYDIA PNEUMONAE ANTIGEN FUSED PROTEIN, DNA CODING FOR THE SAME, RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT CONTAINING THE VECTOR, AND PRODUCTION OF ANTI-CHLAMYDIA PNEUMONAE ANTIBODY
 JOURNAL Patent: JP 1997009976-A 2 14-JAN-1997;

COMMENT HITACHI CHEM CO LTD
 OS None
 OC Artificial sequences.
 PN JP 1997009976-A/2
 PD 14-JAN-1997
 PR 14-MAR-1996 JP 1996057409
 PR 28-APR-1995 JP 95P 106006
 PI IZUTSU HIROSHI, OBARA KAZUHIKO, MATSUMOTO AKIRA PC
 C12N15/09, C07H21/04, C07K14/295, C07K14/47, C07K19/00, C12N1/21, PC
 C12P21/02,
 PC G01N33/569, G01N33/571//A61K39/395, A61K49/00, C12P21/08, PC
 (C12N15/09, C12R1:01),
 PC (C12N1/21, C12R1:19), (C12P21/02, C12R1:19), (C12P21/02, C12R1:91),
 PC (C12P21/08),
 PC (C12R1:91),
 CC strandedness: Double;
 CC topology: Linear;
 FH Key
 FH source
 FT 1. .1296
 FT /organism='Artificial sequences' FT
 mat_peptide 1. .1296
 FT /product='fusion protein of DHFR and antigen peptide'
 FT 1. .480
 FT misc_feature
 FT /note='DHFR coding region'
 FT misc_feature 484..1296
 FT /note='antigen peptide coding region'.
 FEATURES
 source
 1. .1296
 /organism='unidentified'
 /db_xref='taxon:32644'
 BASE COUNT 407 a 299 c 325 g 265 t
 ORIGIN

Alignment Scores:
 Pred. No.: 2.86e-59 Length: 1296
 Score: 1245.00 Matches: 259
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 52.55% Indels: 0
 DB: 6 Gaps: 0

US-09-889-314-2 (1-496) x E12526 (1-1296)

QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23
 DB 484 ATGCTATTTCATCTTCTTCTGAGCCTGACAAATCAAAAAATATCATGCTCAAGTTCTGT 543

QY 24 ThrSerThrProGlnGlnValProGlnGlnAspLysLeuSerGlyAsnGlnIleThrLysGln 43
 DB 544 ACATCGACACCCAGGCGGTGCCCCAACAGATAGCTGTGTGCGAACGAAACGACAA 603

QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63
 DB 604 ATACAGCAACACGCTGAGGTGTAACAAACAGTGAAGGCGATGCCACATATGCGGT 663

QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83
 DB 664 GCTTCTGAAAAGACAAATCTCTCGACTACAAAAACAGAACGCTCCACAAACAGGGA 723

QY 84 ValAlaIleAlaGlyLysGlnSerSerSerGlnLysAlaGlyAlaAspThrGlnValSer 103
 DB 724 GTTCTGCTGGGAAAGATCTCTCAAGATCAAAAGCGAGTGTGATCTGAGATATCA 783

QY 104 G1ValAlaIleAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
 DB 784 GAGCGGCTGCTACTACAGCATCAATACATGCAACAAAANTGCTATGCAACCTCTATT 843

QY 124 GlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143
 DB 844 GAAGAGCGGACCAAAAGTATGAGTACCTTACAGTCACTTCAAAACCTCAGTGCAGCG 903

OY 184 GIUIEGLYLEUALALEUALALYSALAIIEGINTHrleuglyGluAlaThrIysSerAla 203
|||||
DB 541 GAATCGGACCTGCGCTTGCTTAAGCAATTCAGACATTGGGAAAGCCCAAAATCTGCC 600
OY 204 LeuSerAsnTYrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlu 223
|||||
DB 601 TTATCTACTATGCAAGTAGACAGACAGACAGCAACAAATAAATACTAGCTTAGAA 660
OY 224 LysGlnAlaIleLysIleAspLysGluArgGluGluTYrGlnGlnImetLysAlaIleGlu 243
|||||
DB 661 AAGCAAGCGATTAATAATGATTAAGACAGAGAAAGATACCAAGATGAAGCTCGCCGAA 720
OY 244 GlnLysSerLysAspLeuGlnLysIleThrMetAspThrValAsnThrValMetIleAla 262
|||||
DB 721 CAGAAGTCTAAAGATCTCGAAGGAACATGATGATCTGTAATCTGTGATGATCGCG 777
RESULT 33
E12542 813 bp DNA linear PAT 27-APR-1998
LOCUS E12542
DEFINITION DNA encoding an antigen peptide of Chlamydia pneumoniae.
ACCESSION E12542.1 GI:3251375
VERSION E12542.1
KEYWORDS JP 1997015244-A/2.
SOURCE unclassified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 813)
AUTHORS Izutsu, H. and Matsumoto, A.
TITLE MEASURING METHOD FOR ANTICHLAMYDIA-PNEUMONIA ANTIBODY AND REAGENT
THEREFOR AS WELL AS DIAGNOSTIC MEDICINE FOR CHLAMYDIA PNEUMONIA
INFECTION
JOURNAL Patent: JP 1997015244-A 2 17-JAN-1997;
HITACHI CHEM CO LTD
COMMENT OS Chlamydia pneumoniae
PN JP 1997015244-A/2
PD 17-JAN-1997
PF 14-MAR-1996 JP 1996057410
PR 28-APR-1995 JP 95P 106011
PI IZUTSU HITROSHI, MATSUMOTO AKIRA
PC G01N33/571, C07K14/295, C07K16/12, G01N33/53,
G01N33/569//A61K49/00, C12N15/09;
CC strandedness: Double;
CC topology: Linear;
FH key Location/Qualifiers
FT source 1. 813
FT mat_peptide 1. 813
FT /organism='Chlamydia pneumoniae' strain='YK41'
FT /product='antigen peptide'.
FEATURES
SOURCE 1. 813
/organism="unclassified"
/db_xref="taxon:32644"
BASE COUNT 286 a 181 c 187 g 159 t
ORIGIN
Alignment Scores:
Pred. No.: 1.63e-59 Length: 813
Score: 1245.00 Matches: 259
Percent Similarity: 100.00% Conserved: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 52.55% Indels: 0
Gaps: 0
US-09-889-314-2 (1-496) x E12542 (1-813)

DB 61 ACATGCACACCCAGGCGGTGCCCAACAGATTAAGTGTCTGCAAGCAAGCAAA 120
OY 44 IIEGInGInThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThrIleAlaGly 63
|||||
DB 121 ATACAGCAAAACGTCAGAGGTAAAAACCTGAGATGGAAGGAGATCCATTTCTGGT 180
OY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGly 83
|||||
DB 181 GCTTCTGCAAAAGACAAAACCTCTCGACTACAAAAACGAAACAGCTCCACAAAGGGA 240
OY 84 ValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyValSer 103
|||||
DB 241 GTTGCTGTGGGAAAGAAATCCACAAAAGTCAAAGGACAGGTGCTGATCTGAGTATCA 300
OY 104 GYAlaAlaAlaIleThrThrAlaSerAsnThrAlaThrLysIleAlaImeGlnThrSerIle 123
|||||
DB 301 GGAGCGGCTGCTACTACACATCAATACGTGCAACAAAATTGCTATGACAGCTCTATT 360
OY 124 GlnGluAlaSerLysSerMetGluSerThrIleGlnGluSerLysLeuSerAlaAla 143
|||||
DB 361 GAAGAGCGCAGCAAAAGTATGAGTCTACCTTAGAGTCACTTCAAAAGCTCAGTCCGCG 420
OY 144 GlnMetLysGluValGlnAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlySer 163
|||||
DB 421 CAATGAAAGAAAGTGCAGAGCGGTGTGTGCTGCTCCCTCAGGGAAGAAAGTTCCGGTTCC 480
OY 164 AlaLysLeuGluThrProGluLeuProLysProGlyValThrProArgSerGluValIle 183
|||||
DB 481 GCAAAATTGGAAACACCTGAGCTCCCAAGCCGGGGTGCACCAAGATCAAGAGTTATAC 540
OY 184 GIUIEGLYLEUALALEUALALYSALAIIEGINTHrleuglyGluAlaThrIysSerAla 203
|||||
DB 541 GAATCGGACCTGCGCTTGCTTAAGCAATTCAGACATTGGGAAAGCCCAAAATCTGCC 600
OY 204 LeuSerAsnTYrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlu 223
|||||
DB 601 TTATCTACTATGCAAGTAGACAGACAGACAGCAACAAATAAATACTAGCTTAGAA 660
OY 224 LysGlnAlaIleLysIleAspLysGluArgGluGluTYrGlnGlnImetLysAlaIleGlu 243
|||||
DB 661 AAGCAAGCGATTAATAATGATTAAGACAGAGAAAGATACCAAGATGAAGCTCGCCGAA 720
OY 244 GlnLysSerLysAspLeuGlnLysIleThrMetAspThrValAsnThrValMetIleAla 262
|||||
DB 721 CAGAAGTCTAAAGATCTCGAAGGAACATGATGATCTGTAATCTGTGATGATCGCG 777
RESULT 34
AR122115 1296 bp DNA linear PAT 16-MAY-2001
LOCUS AR122115
DEFINITION Sequence 18 from patent US 6165478.
ACCESSION AR122115
VERSION AR122115.1 GI:14106432
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1296)
AUTHORS Izutsu, H., Obara, K. and Matsumoto, A.
TITLE DNA encoding Chlamydia pneumoniae antigenic polypeptide
JOURNAL Patent: US 6165478-A 18 26-DEC-2000;
FEATURES
SOURCE 1. 1296
/organism="unknown"
BASE COUNT 407 a 299 c 325 g 265 t
ORIGIN
Alignment Scores:
Pred. No.: 2.86e-59 Length: 1296
Score: 1245.00 Matches: 259
Percent Similarity: 100.00% Conserved: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 52.55% Indels: 0
Gaps: 0

Query Match: 52.55% Indels: 0
 DB: 6 Gaps: 0
 US-09-889-314-2 (1-496) x AR122106 (1-813)

QY 4 MetSerIleSerSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23
 |||||
 Db 1 ATGTCTATTTCATCTTCTTCAGACCTGACATCAAAAATATCATGCTCAAGTCTCG 60
 QY 24 ThrSerThrProGlnGlnValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln 43
 |||||
 Db 61 ACATGCACACCCACGCGCGCCCAACAGATAGCTGCTGCAACGAAACAGCAAA 120
 QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63
 |||||
 Db 121 ATACAGCAACACGCTACGAGGTAAACACTGACATGGAAGCCATGCTATTGCTGCT 180
 QY 64 AlaSerGlyLysAspLysThrSerThrThrLysThrGlnThrAlaProGlnGlnGly 83
 |||||
 Db 181 GCTTGTGAAAAGCAAAACTTCTCTGACTACAAAACAAAAGAGCTCCACAAACAGGA 240
 QY 84 ValAlaAlaGlySerGlySerSerGlySerGlyLysAlaGlyAlaAspThrGlyValSer 103
 |||||
 Db 241 GTTGTCTGCGGGAAGATCTCAGAAAGTCAAAAGCAGCTGCTGATCTGAGATATCA 300
 QY 104 GlyAlaAlaAlaThrThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
 |||||
 Db 301 GGAGCGGCTGCTACTACAGATCAAAATACGCAACAAAATTCCTATGACAGCTCTATT 360
 QY 124 GlnGlnAlaSerLysSerSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143
 |||||
 Db 361 GAAGAGCGCGCAAAAGTATGAGAGTCTACTTACAGTCAAAAGCCACAGTCCGCG 420
 QY 144 GlnMetLysGlnValGlnAlaValAlaValAlaAlaLeuSerGlyLysSerGlySer 163
 |||||
 Db 421 CAAATGAAAGAACTCGAAGCGGTGTGTGCTGCCCTCTCAGGAGAAAGTTGCGGTTC 480
 QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlnValThrProArgSerGlnValIle 183
 |||||
 Db 481 GCAAAATTTGAAACCTGAGCTCCCAAGCCCGGGGTGACCCCAAGATCAAGGTTATC 540
 QY 184 GlnIleGlyLeuAlaLeuAlaValAlaIleGlnThrLeuGlnGlnAlaThrLysSerAla 203
 |||||
 Db 541 GAATGTGAGACTGCGCGTGTGTAAGCAATTCAGACATGGGAGAGCCACAAAATCGCC 600
 QY 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGln 223
 |||||
 Db 601 TTATCTAACTATGCAAGTACACAGCAGACAGCAACCAATAAATACTAGGTCTAGAA 660
 QY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 243
 |||||
 Db 661 AACCAAGCGTAAATAATCGTAAAGACGAGAGATACCAAGAGATGAGAGGCTGCCGAA 720
 QY 244 GlnLysSerLysAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 262
 |||||
 Db 721 CAGAAGCTTAAGATCTCGAAGGAACAATGATCTGCTACTGATGATGATCGCG 777

RESULT 32
 E12521 813 bp DNA linear PAT 27-APR-1998
 LOCUS E12521
 DEFINITION DNA encoding an antigen peptide of Chlamydia pneumoniae.
 ACCESSION E12521
 VERSION E12521.1 GI:3251354
 KEYWORDS JP 1997009974-A/2.
 SOURCE unclassified.
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 813)
 AUTHORS Izutsu, H. and Matsumoto, A.
 TITLE ANTIGEN POLYPEPTIDE INVOLVED WITH CHLAMYDIA PNEUMONAE, DNA CODING
 FOR THE SAME, RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT
 CONTAINING THE VECTOR, AND PRODUCTION OF ANTI-CHLAMYDIA PNEUMONAE
 JOURNAL Patent: JP 1997009974-A 2 14-JAN-1997;

COMMENT HITACHI CHEM CO LTD
 OS Chlamydia pneumoniae
 PN JP 1997009974-A/2
 PD 14-JAN-1997
 PF 20-SEP-1995 JP 1995242095
 PR 20-SEP-1994 JP 94P 224711, 28-APR-1995 JP 95P 106010 PI
 IZUTSU HIROSHI, MATSUMOTO AKIRA
 PC C12N15/09, C07H21/04, C07K14/295, C07K16/12, C12N1/21, C12N15/02,
 PC C12P21/08
 PC G01N33/53, G01N33/571//A61K39/118, A61K39/118, A61K39/118, PC
 (C12N1/21, C12R1:19),
 PC (C12P21/08, C12R1:91);
 CC strandedness: Double;
 CC topology: Linear;
 FH key Location/Qualifiers
 FH FT source 1..813
 FT FT /strain='YK41', /organism='Chlamydia pneumoniae' FT
 FT FT mat_peptide 1..813
 FT FT /product='antigen peptide',
 FT source 1..813
 FT FT /organism='unidentified'
 FT FT /db_xref='taxon:32644'

BASE COUNT 286 a 181 c 187 g 159 t
 ORIGIN

Alignment Scores:
 Pred. No.: 1,63e-59 Length: 813
 Score: 1245.00 Matches: 259
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 52.55% Indels: 0
 DB: 6 Gaps: 0

US-09-889-314-2 (1-496) x E12521 (1-813)

QY 4 MetSerIleSerSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23
 |||||
 Db 1 ATGTCTATTTCATCTTCTTCAGACCTGACATCAAAAATATCATGCTCAAGTCTCG 60
 QY 24 ThrSerThrProGlnGlnValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln 43
 |||||
 Db 61 ACATGCACCCACGCGCGCCCAACAGATAGCTGCTGCAACGAAACAGCAAA 120
 QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63
 |||||
 Db 121 ATACAGCAACACGCTACGAGGTAAACACTGACATGGAAGCCATGCTATTGCTGCT 180
 QY 64 AlaSerGlyLysAspLysThrSerThrThrLysThrGlnThrAlaProGlnGlnGly 83
 |||||
 Db 181 GCTTGTGAAAAGCAAAACTTCTCTGACTACAAAACAAAAGAGCTCCACAAACAGGA 240
 QY 84 ValAlaAlaGlyLysSerSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 103
 |||||
 Db 241 GTTGTCTGCGGGAAGATCTCAGAAAGTCAAAAGCAGCTGCTGATCTGAGATATCA 300
 QY 104 GlyAlaAlaAlaThrThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
 |||||
 Db 301 GGAGCGGCTGCTACTACAGATCAAAATACGCAACAAAATTCCTATGACAGCTCTATT 360
 QY 124 GlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143
 |||||
 Db 361 GAAGAGCGCGCAAAAGTATGAGAGTCTACTTACAGTCAAAAGCCACAGTCCGCG 420
 QY 144 GlnMetLysGlnValGlnAlaValAlaValAlaAlaLeuSerGlyLysSerGlySer 163
 |||||
 Db 421 CAAATGAAAGAACTCGAAGCGGTGTGTGCTGCCCTCTCAGGAGAAAGTTGCGGTTC 480
 QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlnValThrProArgSerGlnValIle 183
 |||||
 Db 481 GCAAAATTTGAAACCTGAGCTCCCAAGCCCGGGGTGACCCCAAGATCAAGAGTTATC 540

Db 481 GCAAAATTGGAACACTGAGCTCCCAAGCCGGGGTGACACCAAGATCAGAGTTATC 540
Qy 184 GtUleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGluAlaThrLysSerIa 203
Db 541 GAAATCGAGCTCGCGCTTGAACCAATTCACATTGGGAGAACCAAAATCTGCC 600
Qy 204 LeuSerAsnTYrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlu 223
Db 601 TTATCTAACTATGCAAGTACACACACAGACCAACAAATAAATAGTCTAGAA 660
Qy 224 LysGlnAlaIleLysIleAspLysGluArgGluTYrGlnGluMetLysAlaIaGlu 243
Db 661 AAGCAAGCGATAAATGATTAAGACGAGAAATACCAAGATGAGCGTCCGAA 720
Qy 244 GlnLysSerLysAspLeuGlnGluTYrThrMetAspThrValAsnThrValMetIleAla 262
Db 721 CAGAAGTCTAAGATCTCGAAGAACAAATGATACTGTCAATACTGTGATGATCGG 777
RESULT 30
E12543 777 bp DNA linear PAT 27-APR-1998
LOCUS E12543
DEFINITION DNA encoding an antigen peptide of Chlamydia pneumoniae.
ACCESSION E12543
VERSION E12543.1 GI:3251376
KEYWORDS JP 1997015244-A/3.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 777)
AUTHORS Izutsu,H. and Matsumoto,A.
TITLES MEASURING METHOD FOR ANTICHLAMYDIA-PNEUMONIA ANTIBODY AND REAGENT THEREFOR AS WELL AS DIAGNOSTIC MEDICINE FOR CHLAMYDIA PNEUMONIA INFECTION
JOURNAL Patent: JP 1997015244-A 3 17-JAN-1997;
COMMENT HITACHI CHEM CO LTD
OS Chlamydia pneumoniae
PN JP 1997015244-A/3
PD 17-JAN-1997
PF 14-MAR-1996 JP 1996057410
PR 28-APR-1995 JP 95P 106011
PI IZUTSU HIROSHI, MATSUMOTO AKIRA
PC GOIN33/571,C07K14/295,C07K16/12,G01N33/53,
PC GOIN33/569//A61K49/00,C12N15/09;
CC strandedness: Double;
CC topology: Linear;
FH key Location/Qualifiers
FH source 1..777
FT /strain='YK41', /organism='Chlamydia pneumoniae' FT
FT mat-peptide 1..777
FT Location/Qualifiers
FT 1..777
FT /db_xref='taxon:32644'
FEATURES
source Location/Qualifiers
BASE COUNT 277 a 175 c 176 g 149 t
ORIGIN
Alignment Scores:
Pred. No.: 1.55e-59 Length: 777
Score: 1245.00 Matches: 259
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 52.558 Indels: 0
DB: 6 Gaps: 0
US-09-889-314-2 (1-496) x E12543 (1-777)
Qy 4 MetSerLysSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValIleu 23
Db 1 ATGTCTATTCATCTCTCTTCAGGACCTGACAAATCAAAAAATATCATGCTCAAGTTCTG 60

Qy 24 ThrSerThrProGlnGlyValProGlnLysLysLeuSerGlyAsnGluThrLysGln 43
Db 61 ACATGACACACCCAGAGCGGTGCCCAACAGATAAAGTGTCTGGCAACGAACGAA 120
Qy 44 IleGlnGlnThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThrIleAlaIy 63
Db 121 ATACAGCAACACGTCAGGGTAAACACAGAGGAGGAAGCGAGCCACTATTCCTGT 180
Qy 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGly 83
Db 181 GCTTGTGGAAGAACAAACTTCCTCGACTACAAAAAGAAACGACGCTCCAAACAGGA 240
Qy 84 ValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyValSer 103
Db 241 GTTCTGTCTGGGAAAGAAATCTCAGAAAGTCNAAGGACGAGTGTGATCTGAGTATCA 300
Qy 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGluThrSerIle 123
Db 301 GGAGCGCTGTCTACTACAGCATCAATCTGCACAAAAATGCTATGACAGACTCTATT 360
Qy 124 GluGluAlaSerLysSerMetGluSerThrLeuGlnSerLeuSerAlaAla 143
Db 361 GAAAGAGCGGACGAAAGATATGAGTCTACTAGATCACTTCAAAAGCTCAGTCCGG 420
Qy 144 GlnMetLysGluValGluAlaValAlaAlaAlaLeuSerGlyLysSerSerGlySer 163
Db 421 CAATGAAAGAAAGTGAAGCGGTGTGTGTCTGCCCTCAGGAAAAAGTCCGGGTTCC 480
Qy 164 AlaLysLeuGluThrProGluLeuProLysProGlyValThrProArgSerGluValIle 183
Db 481 GCAAAATTGGAACACCTGAGTCCCAAGCCGGGGGTACACCAAGATCAAGGTTATC 540
Qy 184 GtUleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGluAlaThrLysSerAla 203
Db 541 GAAATCGAGCTCGCGCTTGAACCAATTCAGACATTGGAGAACCAAAATCTGCC 600
Qy 204 LeuSerAsnTYrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlu 223
Db 601 TTATCTAACTATGCAAGTACACACACAGACCAACCAAAATAAATAGTCTAGAA 660
Qy 224 LysGlnAlaIleLysIleAspLysGluArgGluGluTYrGlnGluMetLysAlaIaGlu 243
Db 661 AAGCAAGCGATAAATGATTAAGACGAGAAATATCAAGATGAGTGAAGCTGCCGAA 720
Qy 244 GlnLysSerLysAspLeuGlnGluTYrThrMetAspThrValAsnThrValMetIleAla 262
Db 721 CAGAAGTCTAAGATCTCGAAGAACAAATGATACTGTCAATACTGTGATGATCGG 777
RESULT 31
AR122106 813 bp DNA linear PAT 16-MAY-2001
LOCUS AR122106
DEFINITION Sequence 4 from patent US 6165478.
ACCESSION AR122106
VERSION AR122106.1 GI:14106423
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 813)
AUTHORS Izutsu,H., Obara,K. and Matsumoto,A.
TITLES DNA encoding Chlamydia pneumoniae antigenic polypeptide
JOURNAL Patent: US 6165478-A 4 26-DEC-2000;
FEATURES
source Location/Qualifiers
BASE COUNT 286 a 181 c 187 g 159 t
ORIGIN
Alignment Scores:
Pred. No.: 1.63e-59 Length: 813
Score: 1245.00 Matches: 259
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0

Db 887 GGCTGTAAGAAAGCAGACGATTAATAATCGATTAAGAAAGCAGACAGATTAACCAAGATGTAAG 946

QY 241 ALaAlaGluGlnLysSerLysAspLeuGluGlyThrMetAspThrValaSerThrValMet 260
|||||

Db 947 GCGCGCAACAGAAAGTCTTAAGATCTCGAAGGAGAACATGATGATCTGTCATCTGATGATG 1006

QY 261 lIeAla 262
|||||

Db 1007 ATCGCG 1012

RESULT 27

LOCUS E12265 1048 bp DNA linear PAT 27-APR-1998

DEFINITION DNA encoding an antigen of Chlamydia pneumoniae.

ACCESSION E12265

VERSION E12265.1 GI:3251099

KEYWORDS JP 1996304404-A/2.

SOURCE JP 1996304404-A/2.

ORGANISM unidentified.

REFERENCE 1 (bases 1 to 1048)

AUTHORS Iizutsu,H. and Matsumoto,A.

TITLE DETECTION/MEASUREMENT AND REAGENT FOR ANTICHLADYDIA/ NEUMONIE ANTIBODY AND DIAGNOSIS AGENT OFCHLADYDIA/ NEUMONIE INFECTION

JOURNAL Patent: JP 1996304404-A 2 22-NOV-1996;

COMMENT HITACHI CHEM CO LTD

OS Chlamydia pneumoniae

PN JP 1996304404-A/2

PD 22-NOV-1996

PI 28-APR-1995 JP 1995106014

PI IZUTSU HIROSHI, MATSUMOTO AKIRA

PC G01N33/569, C07K14/295, G01N33/53, G01N33/571//A61K9/00, PC C12M15/09, C12P21/02

PC (C12P21/02,C12R1:19);

CC strandedness: Double;

CC topology: Linear;

CC hypothetical: No;

CC anti-sense: No;

CC fragment_type: N-Terminal Fragment;

FH key Location/Qualifiers

FH source 1. 1048

FT /organism='Chlamydia pneumoniae' FT

FT /strain='YK-41' /clone='53-3S' FT

FT CDS 236..1012

FT /product='53kDa antigen of Chlamydia FT

FT pneumoniae' /partial.

FEATURES

source Location/Qualifiers

1. 1048

BASE COUNT 360 a 241 c 231 g 216 t

ORIGIN

Alignment Scores:

Pred. No.: 2.67e-60 Length: 1048

Score: 1262.00 Matches: 262

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 53.27% Indels: 0

DB: Gaps: 0

US-09-889-314-2 (1-496) x E12265 (1-1048)

QY 1 AspThrAsnMetSerIleSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20
|||||

Db 227 GATACAAACATGCTATTTCATCTCTTCAGACCTGACATCAATAAATAATCATCTCT 286

QY 21 GlnValLeuThrSerThrProGlnGlnValProGlnGlnAspLysLeuSerGlyAsnGlu 40
|||||

Db 287 CAAGTTCGACATCGACACCCAGGCGTGCCCAACAAAGATTAAGCTGTCTGCAACGAA 346

QY 41 ThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThr 60
|||||

Db 347 ACCAGCAATATACAGCAACACGCTCAGGGTAAAAACACTAGATGGAAAGCATGCCACT 406

QY 61 lIeAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnAlaPro 80
|||||

Db 407 ATTGCTGTCTCTTCTGGAAAAAGACAAACCTTCTCGACTCAAAAACAGACAGCTCCA 466

QY 81 GlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThr 100
|||||

Db 467 CAACAGGAGTTCCTGCTGGGAAAGAACTCTCAGAAAGTCAAAAAGCGAGTGTGATACT 526

QY 101 GlnValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120
|||||

Db 527 GGAATACAGAGCGGCTGCTACTACAGCATCAAAATCTCAACAAAATTTGCTATGAC 586

QY 121 ThrSerIleGlnGlnAlaSerLysSerMetGluSerThrLeuGlnSerLeu 140
|||||

Db 587 ACCTCATTAAGAGGCGGACCAAAAGATGAGTACCTTAGAGTCACTTCAAGCCTC 646

QY 141 SerAlaAlaMetLysGluValGluAlaValAlaValAlaLysSerGlyLysSer 160
|||||

Db 647 AGTCCGCGCAAAATGAAGAAGTCGAAGCGGTGTGTTGCTCCCTCCCTCAGGGAAGT 706

QY 161 SerGlySerAlaLysLeuGlnThrProGlnLeuProLysProGlnValThrProArgSer 180
|||||

Db 707 TCGGGTTCGCAAAATGGAAACACCTGAGACCTCCCAAGCCGGGGGAGACCAAGATCA 766

QY 181 GluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlyAlaThr 200
|||||

Db 767 GAGTATATCAAAATCGACACCTGCGCTTGTCTTAAGCATTTGAGAGAGAGCCACA 826

QY 201 LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
|||||

Db 827 AAATCTGCCCTTACTACTATGCAAGTACCAAGCACCAACCAACCAAAATAAATACTA 886

QY 221 GlnLeuGlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnGlnGlnGlnMetLys 240
|||||

Db 887 GGCTGTAAGAAAGCAGATTAATAATCGATTAAGAAAGCAGAAAGATTAAGATGAAG 946

QY 241 ALaAlaGluGlnLysSerLysAspLeuGluGlyThrMetAspThrValaSerThrValMet 260
|||||

Db 947 GCTGCCAACAAGCTTAAGATCTCGAAGGAGAAACATGATGATCTGTCATCTGATG 1006

QY 261 lIeAla 262
|||||

Db 1007 ATCGCG 1012

RESULT 28

LOCUS AR122107 777 bp DNA linear PAT 16-MAY-2001

DEFINITION Sequence 7 from patent US 6165478.

ACCESSION AR122107

VERSION AR122107.1 GI:14106424

KEYWORDS

SOURCE unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 777)

AUTHORS Iizutsu,H., Obara,K. and Matsumoto,A.

TITLE DNA encoding Chlamydia pneumoniae antigenic polypeptide

JOURNAL Patent: US 6165478-A 7 26-DEC-2000;

FEATURES

source Location/Qualifiers

1. 777

BASE COUNT 277 a 175 c 176 g 149 t

ORIGIN

Alignment Scores:

Pred. No.: 1.55e-59 Length: 777

Score: 1245.00 Matches: 259

Percent Similarity: 100.00% Conservative: 0

QY 81 GlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThr 100
DB 467 CAACAGGAGATTGCTGCTGCAAGAAATCCAGAAAGTCAAAAGGAGGTGCTGATATCT 526
QY 101 GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120
DB 527 GGATATACAGAGCGCGCTACTACACATCAATACTGCACAAAAATTGCTATGAG 586
QY 121 ThrSerIleGluGluAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeu 140
DB 587 ACCTTATTGAAGAGGCGAGCAAAAAGTATGAGTCTACCTTAGAGTCACTTCAAAAGCCTC 646
QY 141 SerAlaAlaGlnMetLysGluValGlnAlaValAlaValAlaAlaLeuSerGlyLysSer 160
DB 647 AGTGCCCGGCAAAATGAAAGAGTCAGAACGGTGTGTGTGCTGCCCTCCAGGAAAAAGT 706
QY 161 SerGlySerAlaLysLeuGlnLysThrProGluLeuProLysProGlyValThrProArgSer 180
DB 707 TCGGGTTCGCAAAATTGGAAACACCTGAGCTCCCAAGCCCGGGGTGACACCAAGATCA 766
QY 181 GluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThr 200
DB 767 GAGGTATCGAAATCGAGACTCGCGCTGCTTAAGCAATTGAGACATTTGGAGAACCCACA 826
QY 201 LysSerAlaLeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
DB 827 AAATCTGCTTATCTACTATGCAAGTACCAAGACAGCAAGCAGACCAAAATCAATCACTA 886
QY 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGluArgGluGluTyrGlnGluMetLys 240
DB 887 GGTCTAGAAAAAGCAAGCATATAAATGATTAAGAACAGAAAGATACCAAGATGAG 946
QY 241 AlaAlaGlnGlnLysSerLysAspLeuGlnGluTyrMetAspThrValAsnThrValMet 260
DB 947 GCTGCCGCAAGAGTCTAAGATCTCGAAGGAAACATGATCTGCAATCTGTGATG 1006
QY 261 IleAla 262
DB 1007 ATCGCG 1012
RESULT 26
E11479
LOCUS E11479 1048 bp DNA linear PAT 29-SEP-1997
DEFINITION DNA encoding 53k antigen peptide from Chlamydia neumonie.
ACCESSION E11479
VERSION E11479.1 GI:22025115
KEYWORDS JP 1996143594 -A/2.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1048)
AUTHORS Izutsu,H. and Matsumoto,A.
TITLE ANTIGEN POLYPEPTIDE FOR CHLAMYDIA NEUMONEAE, DNA CODING FOR THE
SAME, RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT
CONTAINING THE RECOMBINANT VECTOR, AND PRODUCTION OF ANTI-CHLAMYDIA
NEUMONEAE
JOURNAL Patent: JP 1996143594 -A 2 04 -JUN-1996;
HITACHI CHEM CO LTD
OS Chlamydia neumonie
PN JP 1996143594 -A/2
PD 04-JUN-1996
PF 28-APR-1995 JP 1995106013
PR 20-SEP-1994 JP 94P 224711
PI IZUTSU HIROSHI, MATSUMOTO AKIRA
PC C07K14/295, C07K16/12, C12N15/02, C12N15/09, C12P21/08//C12N1/21,
PC C12P21/08
PC C12R1:91), (C12N1/21, C12R1:19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH key Location/Qualifiers

FH source 1..1048
FT /organism='Chlamydia neumonie' FT
FT /strain='YK-41' /clone='53-3S'
FT 5'UTR 1..235
FT CDS 236..1012
FT /product='53k antigen peptide from Chlamydia
FT neumonie'
FT 1013..1048.
FEATURES
source 1..1048
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 360 a 241 c 231 g 216 t
ORIGIN
Alignment Scores:
Pred. No.: 2,67e-60 Length: 1048
Score: 1262.00 Matches: 262
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.27% Indels: 0
DB: Gaps: 0
US-09-889-314-2 (1-496) x E11479 (1-1048)
QY 1 AspThrAsnMetSerIleSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20
DB 227 GATACAAACATGCTATTTCATCTTCTTCTGAGACCTGCAATCAAAAAATATCATGTCT 286
QY 21 GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyValGlnGlu 40
DB 287 CAAGTCTGACATCGACACCCAGGGCGTCCCCAACAAAGATATACCTTCTGGCAACCA 346
QY 41 ThrLysGlnIleGlnGlnThrArgGlnGlnLysAsnThrGluMetGluSerAspAlaThr 60
DB 347 ACGAAGCAAAATACACCAAAACAGTCAAGGTAAAAACACTGAGATGAAAGCATCCACT 406
QY 61 IleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaPro 80
DB 407 ATTGCTGTGCTCTGGAAGAACAAACTTCTCTGACTACAAAAACAGAAACACTCTCA 466
QY 81 GlnGlnGlyValAlaAlaGlyLysGluSerSerGlnLysAlaGlyAlaAspThr 100
DB 467 CAACAGGAGATTGCTGCTGCAAGAAATCCAGAAAGTCAAAAGGAGGTGCTGATATCT 526
QY 101 GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120
DB 527 GGATATACAGAGCGCGCTACTACACATCAATACTGCACAAAAATTGCTATGAG 586
QY 121 ThrSerIleGluGluAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeu 140
DB 587 ACCTTATTGAAGAGGCGAGCAAAAAGTATGAGTCTACCTTAGAGTCACTTCAAAAGCCTC 646
QY 141 SerAlaAlaGlnMetLysGluValGlnAlaValAlaValAlaAlaLeuSerGlyLysSer 160
DB 647 AGTGCCCGGCAAAATGAAAGAGTCAGAACGGTGTGTGTGCTGCCCTCCAGGAAAAAGT 706
QY 161 SerGlySerAlaLysLeuGlnLysThrProGluLeuProLysProGlyValThrProArgSer 180
DB 707 TCGGGTTCGCAAAATTGGAAACACCTGAGCTCCCAAGCCCGGGGTGACACCAAGATCA 766
QY 181 GluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThr 200
DB 767 GAGGTATCGAAATCGAGACTCGCGCTGCTTAAGCAATTGAGACATTTGGAGAACCCACA 826
QY 201 LysSerAlaLeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
DB 827 AAATCTGCTTATCTACTATGCAAGTACCAAGACAGCAAGCAGACCAAAATCAATCACTA 886
QY 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGluArgGluGluTyrGlnGluMetLys 240

```

OY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGlnGly 83
    |||||
Db 720 GCTTGGAAAAAGACAAACTCTCGACTACAAAAACGAAACAGCTCCACACAGGGA 779
OY 84 ValAlaIaIaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyValSer 103
    |||||
Db 780 GTTGGCTGGGGAAGAACTCTCAGAAAGTCAAAAGCGAGTGTGATCTGAGATCA 839
OY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
    |||||
Db 840 GGAGGGCTGCTACTACACATCAATACTGCACAAAAATTCATGACACTCTATT 899
OY 124 GlnGluAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeuSerAlaIa 143
    |||||
Db 900 GAGAGCGGACGCAAAAGTATGAGTCTACCTTACGATCTTCAAAAGCTCAGTGGCGG 959
OY 144 GlnMetLysGluValGluAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlySer 163
    |||||
Db 960 CAAATGAAAGAGTGAAGCGGTTGTTGCTGCCCTCAGGCAAAAGTTCGGGTTCC 1019
OY 164 AlaLysLeuGluThrProGluLeuProLysProGlyValThrProArgSerGlyValIle 183
    |||||
Db 1020 GCAAAATTTGAAACACCTGAGTCCCAAGCCCGGGTGACACCAAGATCAGAGTTATC 1079
OY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThrLysSerAla 203
    |||||
Db 1080 GAAATCGGACTCGCGCTTCTAAAGCAATTCAGACTTGGAGAAACCCAAATATGCCC 1139
OY 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlu 223
    |||||
Db 1140 TTAATCAATATGCAAGTACACCAAGCAGACCAAGCAAAATTAAGTAGTATAGAA 1199
OY 224 LysGlnAlaIleLysIleAspLysGluArgGluGluThrGlnGluMetLysAlaAlaGlu 243
    |||||
Db 1200 AAGCAGCGCATAAATCATTAAGAACGAGAAAGATACCAGATGAGTGAAGCTGCGGAA 1259
OY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaVal 263
    |||||
Db 1260 CAGAACTCTAAAGATCTCAGAGCAACATGATCTGCTATCTGTGATGCGGAA- 1318
OY 264 SerValAlaIleThrVal-----IleSerIleValAla 274
    |||||
Db 1319 GGGTTCGAAATTGCATGGGGCCCTTAATTAATTAATTCGAGAGATCCAGATCTAATCG 1378
OY 275 AlaIlePheThrCysGlyAlaGlyLeu-----AlaAlaGlyAlaAlaVal 292
    |||||
Db 1379 ATGATCTCTACGCGGAGCATCTGTGGCGGCATCCGCGCCACAGGTGGGTTGCT 1438
OY 293 GlyAlaAla----- 295
    |||||
Db 1439 GCGGCTTA-TATCGCCGACATCAGCGATGGGAGATCGGGCTCGCCACTTGGGCTCAT 1497
OY 296 -----AlaAlaGlyGlyAlaAlaGlyAla 303
    |||||
Db 1498 GAGCGCTGTTCGGCGGTGATGGTGCAGGCGCGTGGCGGGGAGCTGTGGCGCC 1557
OY 304 -----AlaAlaAla----- 306
    |||||
Db 1558 ATCTCCTTCATGACACATCTCTTGGCGGGGCTCTCAACGGCTCAACCTACTACTG 1617
OY 306 ----- 306
    |||||
Db 1618 GGCTGCTTCTTAATGAGAGTGCATTAAGGAGAGCGTCGACCGATGCCCTTGAAGCC 1677
OY 307 -----ThrThrValAlaThrGlnIle 313
    |||||
Db 1678 TTCAACCCAGTACGTCTTCGGTGGCGGGGAGTATGCTGCGCGCACTTATAG 1737
OY 314 ThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValArgGlnAla 333
    |||||
Db 1738 ACTGCTCTTATATCATGCAACTCTGAGAGAG----- 1770

```

```

OY 334 IleThrAlaAlaIle-----LysAlaAlaValLys 343
    |||||
Db 1771 GTGGCGGAGCGGCTCTGGGTCAATTTTCGGCGAGACCGCTTTCGCGAGCGCAGATG 1830
OY 344 SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363
    |||||
Db 1831 ATCGGCTTCCTCCCTTCGATATTCGGAATCTTGACAGCCCTCGCTCAACGCTTCGACT 1890
OY 364 GlyIleSerValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383
    |||||
Db 1891 GGTCCCGCC-----ACCAACGTTTCGGCGAGAG 1920
OY 384 SerLysValIleSerSerLeu----- 390
    |||||
Db 1921 CAGGCAATATGCGCGCATGGCGGCGGCGGCTGAGCTTGTGCTGGCGTTGCGG 1980
OY 391 ThrSerLysTrpValThrValGlyValGlyValAlaAlaAlaProAlaLeuGlyLys 410
    |||||
Db 1981 ACGCGAGCGTGGATGGCTTCCCATTTATGATTTCTGCTCGCGCGCATCGGGATG 2040
OY 411 GlyIleMetGlnMetGlnLeuSerGlnMetGlnAsnValAlaGlnPheGlnLysGlu 430
    |||||
Db 2041 CCGCGCTTCAGGCCCATGCTGTCCAGCGCAGTATGATGACCAATCAGGACAGCTTCAA 2100
OY 431 ValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMet 443
    |||||
Db 2101 ----GATTCGCTCGCGGCTCTTAACACCTTACTGATC 2136
    |||||

RESULT 25
ARI22109 1048 bp DNA linear PAT 16-MAY-2001
LOCUS ARI22109
DEFINITION Sequence 9 from patent US 6165478.
ACCESSION ARI22109
VERSION ARI22109.1 GI:14106426
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1048)
AUTHORS Izutsu,H., Obara,K. and Matsumoto,A.
TITLE DNA encoding Chlamydia pneumoniae antigenic polypeptide
JOURNAL Patent: US 6165478-A 9 26-DEC-2000;
FEATURES
source 1..1048
BASE COUNT 360 a 241 c 231 g 216 t
ORIGIN
Alignment Scores:
Pred. No.: 2,67e-60 Length: 1048
Score: 1262.00 Matches: 262
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.27% Indels: 0
DB: Gaps: 0
US-09-889-314-2 (1-496) x ARI22109 (1-1048)
OY 1 AspThrAsnMetSerIleSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20
    |||||
Db 227 GATACAAACATGCTCTTATCATCTCTTCAGACCTGACATCAAAAAATATCATGCTCT 286
OY 21 GlnValLeuThrSerThrProGlnGlyValAlaProGlnGlnAspLysLeuSerGlyAsnGlu 40
    |||||
Db 287 CAAGTTCGACATCGACACCCCGAGGCGGTGCCCAACAAGATTAAGCTGTGGCAAGAA 346
OY 41 ThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThr 60
    |||||
Db 347 ACGAACAATACAGCAAAACACTGACGGTAAACACATGAGATGGAAGGCAATGCGACT 406
OY 61 IleAlaGlyAlaSerGlyLysAspLysThrSerSerThrLysThrGluThrAlaPro 80
    |||||
Db 407 ATTGCTGTGCTTCTGGAAGAACAACTCTCTCGACTTACAAACACAGAAACAGCTCCA 466

```

Db 960 CAATGAAGAAGTCGAAGGGGTGTGTGTCGCCCTCAGGAGAAAGTTCGGGTCC 1019
QY 164 AAlaIleuGIuThrProGIuLeuProIysProGIyValIThrProArgSerGIuValIle 183
Db 1020 GCAAAATGGAAACCTGAGCTCCCAAGCCCGGGGTGACACCAGATCAGAGATTATC 1079
QY 184 GIuIleGIuLeuAlaIleuAlaIyAlaIleGIuThrIleuGIyGIuIleThrIySerAla 203
Db 1080 GAATTCGAGCTCGGCTTGCTTAAAGCAATTCAGCATTTGGAGAGCCACAAAATCTGCC 1139
QY 204 LeuSerAntIyAlaSerThrGIuAlaGIuAlaAspGIuThrAsnIyLeuGIu 223
Db 1140 TTATCTAATCATGCAATACAGCAGCAAGCAGACCAACAATTAAGTCTGAA 1199
QY 224 LysGIuAlaIleIySIleAspIySGIuArgGIuIuTyrgIuGIuMetIyAlaGIu 243
Db 1200 AAGCAACGATAAATAATGATTAAGAACGAGAAAGATACCAAGATGAAGGCTCCGAA 1259
QY 244 GIuIySerIyAspLeuGIuGIuIyThrMetAspThrValAsnThrValMetIleAlaVal 263
Db 1260 CAGAAAGCTAAAGATCTCGAAGCAAAATGATCTCAATCTGATGATGCGAA- 1318
QY 264 SerValAlaIleThrVal-----IleSerIleValAla 274
Db 1319 GGGTTGCAATTCGATGGGGCCCTTAATTAATTAATCAGAGATCCAGATCTAATCG 1378
QY 275 AAlaIleThrIySGIyAlaGIuAlaGIuLeu-----AlaAlaGIyAlaAlaVal 292
Db 1379 ATGATCTCTACGCGGAGCAGCATGTGGCGGCATCCGCGCAGACGATGCGGTGCT 1438
QY 293 GIyAlaIle----- 295
Db 1439 GGGCCCTTA-TATGCCGACATCCAGATGGGAAGATCGGGCTCGCCACTTCGGGCTCAT 1497
QY 296 -----AlaAlaGIyAlaAlaGIyAla 303
Db 1498 GAGGCGTTGTTGGCGGTATGTGGCGGCGGCCGCGGGGAGCTTTGGGGCC 1557
QY 304 -----AlaAlaIle----- 306
Db 1558 ATCTCTTGATGCACCATTCCTTCGCGCGGGTGTCAAGGCGCTCACTACTGTG 1617
QY 306 ----- 306
Db 1618 GCGCTTCTTAATGCAGAGAGTCCATTAAGGAGAGCGTCGACCGATGCCCTTGAGAGCC 1677
QY 307 -----ThrThrValAlaThrGIuIle 313
Db 1678 TTCAACCAAGTCAGCTCCTTCGCGTGGGCGGCGGAGCATGACTATTCGCGCACTTATG 1737
QY 314 ThrValGIuAlaValGIuAlaValIySGIuAlaValIleThrAlaValArgGIuAla 333
Db 1738 ACTGTCTTCTTATCATCATCACTGATAGAGCAG----- 1770
QY 334 IleThrAlaIle-----LysAlaAlaValIy 343
Db 1771 GTGGCGGAGGCTCTGGGTATTTTGGCGAGACCGCTTCGCTGAGGCGCAGCATG 1830
QY 344 SerGIyIleIyAlaIleIyThrIleValIyAlaIleAlaIyAlaIleSerIy 363
Db 1831 ATGGCGCTGTGGCTTGGGAATCTTGACAGCGCTCGCTCAAGCTTGCTCACT 1890
QY 364 GIyIleSerIyValIleAlaIySGIyThrGIuMetIleAlaIyAsnPhrProIyLeu 383
Db 1891 GGTCCCGCC-----ACCAAAAGTTTCGGCGAGAAAG 1920
QY 384 SerIyValIleSerSerLeu----- 390
Db 1921 CAGGCAATATTCGCGGATCGGCGAGCGCGCTGCGCTGCTGCGCTTCGCG 1980
QY 391 ThrSerIyThrProIyAlaGIyAlaGIyValGIyAlaValIaIleAlaIleGIuIy 410
Db 1981 ACGGAGGCTGTGATGGCTTCCCATTTATGATTTCTTCGGTTCGGGGCATCGGATG 2040

QY 411 GIyIleMetGIuMetGIuLeuSerGIuMetGIuGIuAsnValAlaGIuIleGIuIySGIu 430
Db 2041 CCCGCTTGGAGCCATGCTGTCCAGGACAGATGACATGACATGAGGACAGCTTCAA 2100
QY 431 ValGIyIleuGIuAlaIleAlaIleAspMetIleSerMet 443
Db 2101 ---GGATCGCTCGGGGCTTACCAAGCTTAATCTGCATC 2136
RESULT 24
E12540
LOCUS E12540 5438 bp DNA linear PAT 27-APR-1998
DEFINITION Nucleotide sequence of pCPN53T that contains Chlamydia pneumoniae antigen gene.
ACCESSION E12540
VERSION E12540.1 GI:3251373
KEYWORDS JP 1997015243-A/3.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 5438)
AUTHORS Iizusu,H., Obara,K. and Matsumoto,A.
TITLE MEASURING METHOD FOR ANTI-CHLAMYDIA PNEUMONIA ANTIBODY AND REAGENT THEREFOR AS WELL AS DIAGNOSTIC MEDICINE FOR CHLAMYDIA PNEUMONIA INFECTION
JOURNAL Patent: JP 1997015243-A 3 17-JAN-1997;
HITACHI CHEM CO LTD
OS None
OC Artificial sequences.
PN JP 1997015243-A/3
PD 17-JAN-1997
PF 15-MAR-1996 JP 1996058609
PR 28-APR-1995 JP 95P 106009
PI IZUTSU HIROSHI, OBARA KAZUHIKO, MATSUMOTO AKIRA PC
GO1N33/569, C12N1/21, C12N15/02, C12P1/02, C12P21/08, PC
GO1N33/53,
PC GO1N33/531, GO1N33/571//A61K49/00, C07K14/245, C07K14/39, PC
C07K16/12, C07K19/00,
PC (C12N1/21, C12N1.19), (C12N15/09, C12R1.01), (C12P21/02, C12R1.19);
CC strandedness: Double;
CC topology: Circular;
FH Key Location/Qualifiers
FH FT source 1..5438
FT FT Location/Qualifiers
FEATURES
source 1..5438 /organism='Artificial sequences'.
BASE COUNT 1372 a 1424 c 1391 g 1251 t
ORIGIN
Alignment Scores:
Pred. No.: 1.33e-59 Length: 5438
Score: 1265.00 Matches: 312
Percent Similarity: 62.39% Conservative: 33
Best Local Similarity: 56.42% Mismatches: 75
Query Match: 53.40% Indels: 135
Gaps: 10
US-09-889-314-2 (1-496) x E12540 (1-5438)
QY 4 MetSerIleSerSerSerSerGIyProAspAsnGIuIyAsnIleMetSerGIuValIleu 23
Db 540 ATGCTATTTATCTTCTTCAGGACCTGACATCAAAAAATATCATGTCTCAAGTTCTG 599
QY 24 ThrSerThrProGIuGIuValProGIuIleAspIySGIyAsnGIyThrIySGIu 43
Db 600 ACATGACACCCCAAGGCTGCCCCCAACAGATTAAGCTGTGCGACGAAACGAGCAA 659
QY 44 IleGIuIleThrArgGIuGIuIyAsnThrGIuMetGIuSerAspAlaIleAlaGIy 63
Db 660 ATACGCAAAACAGTCAGGATAAAAACACTGAGATGGAAGGATGCCACTATTGTGTGT 719


```
QY 264 SerValAlaIleThrVal-----IleSerIleValAla 274
      ||| |||
      ||| |||
      ||| |||
Db 1319 GGGGTTGCAATTGGCATGGGGCCCTTAATTAATTAATGAGATGAGATCATGATCTAATG 1378
      ||| |||
      ||| |||
      ||| |||
QY 275 AlaIlePheThrCysGlyAlaGlyLeu-----AlaAlaGlyAlaAlaVal 292
      ||| |||
      ||| |||
      ||| |||
Db 1379 ATGATTCCTTACGCGGAGCATGCGCGGATCAGCGGCGCCACAGGTGGGTTGCT 1438
      ||| |||
      ||| |||
      ||| |||
QY 293 GlyAlaAla----- 295
      ||| |||
      ||| |||
      ||| |||
Db 1439 GCGCGCTA-TATCGCGCATGACCGATGGGAAGATCGGGCTCGCCACTTCGGGCTCAT 1497
      ||| |||
      ||| |||
      ||| |||
QY 296 -----AlaAlaGlyAlaAlaGlyAla 303
      ||| |||
      ||| |||
      ||| |||
Db 1498 GAGCGCTTGTTCGGCGGTATGGTGGCAGCGCCGTGGCGGGGACTGTGGCGCGC 1557
      ||| |||
      ||| |||
      ||| |||
QY 304 -----AlaAlaAla----- 306
      ||| |||
      ||| |||
      ||| |||
Db 1558 ATCTCCTTCATGACACCATCTCTGGCGGCGGCTGCTCAAGCGCTCAACCTACTACTG 1617
      ||| |||
      ||| |||
      ||| |||
QY 306 ----- 306
      ||| |||
      ||| |||
      ||| |||
Db 1618 GGCTGCTTCTTAATGAGAGATGCGCATAAAGGAGAGCTGACCGATGCCCTTGAGAGCC 1677
      ||| |||
      ||| |||
      ||| |||
QY 307 -----ThrValAlaThrGlnIle 313
      ||| |||
      ||| |||
      ||| |||
Db 1678 TTCAACCCAGTACGTCCTTCGCGGTGGCGCGGGCGGCGGATGACTATGCTCGCGGACTTATG 1737
      ||| |||
      ||| |||
      ||| |||
QY 314 ThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValArgGlnAla 333
      ||| |||
      ||| |||
      ||| |||
Db 1738 ACTGTCCTTCTTATCATGCACTGACCTGAGAGACAG----- 1770
      ||| |||
      ||| |||
      ||| |||
QY 334 IleThrAlaAlaIle-----LysAlaAlaValLys 343
      ||| |||
      ||| |||
      ||| |||
Db 1771 GTGCCGCGACCGCTCGGTCATTTTCGGCGAGAGACCGCTTTCGCTGGAGCGGAGCATG 1830
      ||| |||
      ||| |||
      ||| |||
QY 344 SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363
      ||| |||
      ||| |||
      ||| |||
Db 1831 ATCGGCTGTCGCTGCGTATATTCGGAATCTTCGACAGCCCTCGCTCAAGCTTCGCACT 1890
      ||| |||
      ||| |||
      ||| |||
QY 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383
      ||| |||
      ||| |||
      ||| |||
Db 1891 GGTTCGCGC-----ACCAACGCTTTCGCGGAGAG 1920
      ||| |||
      ||| |||
      ||| |||
QY 384 SerLysValIleSerSerLeu----- 390
      ||| |||
      ||| |||
      ||| |||
Db 1921 CAGGCCATTATGCGCGCATGGCGCGCGACGCGCTGGCTACGCTTTCGCTGGCGTTCCGG 1980
      ||| |||
      ||| |||
      ||| |||
QY 391 ThrSerLysTrpValThrValGlyValGlyValAlaAlaAlaProAlaLeuGlyLys 410
      ||| |||
      ||| |||
      ||| |||
Db 1981 ACGCGAGCTGGATGGCTTCGCCCATTTATGATTCCTTCGCTTCGCGGCGCATCGGAGATG 2040
      ||| |||
      ||| |||
      ||| |||
QY 411 GlyIleMetGlnMetGlnLeuSerGlnMetGlnGlnAsnValAlaGlnPheGlnLysGlu 430
      ||| |||
      ||| |||
      ||| |||
Db 2041 CCCGCGTTCAGAGCCATGCTGTCTCAGCAGAGTAGATGACGACCATCAGGAGAGCTTCAA 2100
      ||| |||
      ||| |||
      ||| |||
QY 431 ValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMet 443
      ||| |||
      ||| |||
      ||| |||
Db 2101 ----GATTCGCTCGCGGCTCTTACACGCTTAACCTTCGATC 2136
      ||| |||
      ||| |||
      ||| |||
```

```
RESULT 23
E12527 5438 bp DNA linear PAT 27-APR-1998
LOCUS E12527
DEFINITION Nucleotide sequence of pcPM533r that contains Chlamydia pneumoniae
antigen gene.
ACCESSION E12527
VERSION E12527.1 GI:3251360
KEYWORDS JP 1997009976-A/3.
SOURCE unclassified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 5438)
AUTHORS Izutsu,H., Obara,K. and Matsumoto,A.
```

```
TITLE DIHYDROFOLIC ACID REDUCTASE-CHLAMYDIA PNEUMONAE ANTIGEN FUSED
PROTEIN, DNA CODING FOR THE SAME, RECOMBINANT VECTOR CONTAINING THE
JOURNAL DNA, TRANSFORMANT CONTAINING THE VECTOR, AND PRODUCTION OF
ANTI-CHLAMYDIA PNEUMONAE ANTIBODY
Patent: JP 1997009976-A 3 14-JAN-1997;
HITACHI CHEM CO LTD
COMMENT OS None
OC Artificial sequences.
PN JP 1997009976-A/3
PD 14-JAN-1997
PE 14-MAR-1996 JP 1996057409
PR 28-APR-1995 JP 95P 106006
PI IZUTSU HIROSHI, OBARA KAZUHIKO, MATSUMOTO AKIRA PC
C12N15/09, C07H21/04, C07K14/295, C07K14/47, C07K19/00, C12N1/21, PC
C12P21/02,
PC G01N33/569, G01N33/571//A61K39/395, A61K49/00, C12P21/08, PC
(C12N15/09, C12R1:01),
PC (C12N1/21, C12R1:19), (C12P21/02, C12R1:19), (C12P21/02, C12R1:91),
PC (C12P21/08,
PC (C12R1:91);
CC CC strandedness: Double;
CC topology: Circular;
FH FH Key
FT FT source 1. .5438
FEATURES /organism='Artificial sequences'.
source 1. .5438
BASE COUNT 1372 a 1424 c 1391 g 1251 t
ORIGIN
Alignment Scores:
Pred. No.: 1.33e-59 Length: 5438
Score: 1265.00 Matches: 312
Percent Similarity: 62.39% Conservative: 33
Best Local Similarity: 56.42% Mismatches: 75
Query Match: 53.40% Indels: 135
DB: Gaps: 10
US-09-889-314-2 (1-496) x E12527 (1-5438)
QY 4 MetSerIleSerSerSerGlyProAspGlnLysAsnIleMetSerGlnValLeu 23
      ||| |||
      ||| |||
      ||| |||
Db 540 ATGTCATTTCATCTTCTTCAGGACCTGACATCAAAAATAATCATGTCTCAAGTTCTG 599
      ||| |||
      ||| |||
      ||| |||
QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGluThrLysGln 43
      ||| |||
      ||| |||
      ||| |||
Db 600 ACATCGACACCCAGGCGCGGCCCAACAAGATTAACCTGTCTGCGCAACGAAAGCA 659
      ||| |||
      ||| |||
      ||| |||
QY 44 TleGlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63
      ||| |||
      ||| |||
      ||| |||
Db 660 ATACAGCAAAACAGCTAGGCTAAACACTGAGATGGAAGCGATGCCACTATTCCTGGT 719
      ||| |||
      ||| |||
      ||| |||
QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGly 83
      ||| |||
      ||| |||
      ||| |||
Db 720 GCTTCGGAAGAACAAACTTCCTCGACTACAAAACGAAACAGCTCCACACAGGGA 779
      ||| |||
      ||| |||
      ||| |||
QY 84 ValAlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThrGlyValSer 103
      ||| |||
      ||| |||
      ||| |||
Db 780 GTTGCTGTGTGGAAAGAAATCTCAGAAAGTCAAAAAGCAGGTGCTGATCTGAGATCA 839
      ||| |||
      ||| |||
      ||| |||
QY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
      ||| |||
      ||| |||
      ||| |||
Db 840 GGAGCGGCTGTACTACACATCAAAATTAATGCAACAAAATTTGCTATGAGACCTCTATT 899
      ||| |||
      ||| |||
      ||| |||
QY 124 GluGluAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143
      ||| |||
      ||| |||
      ||| |||
Db 900 GAAAGAGCGAGCAAAAGTATGGAGTCTACCTTAGAATCACTTCAAGGCTCAGTGGCGG 959
      ||| |||
      ||| |||
      ||| |||
QY 144 GlnMetLysGluValGlnAlaValValAlaAlaLeuSerGlyLysSerSerGlySer 163
      ||| |||
      ||| |||
      ||| |||
```


QY 296 -----AlaAlaGlyGlyAla 300
DB 1709 GCGGCTCATGAGCGCTGTTGCGGGTGGAATGGTGGAGGCCGCTGGCGGGGAGCTG 1768
QY 301 AlaGlyAla-----AlaAlaAla----- 306
DB 1769 TTGGGCGCCATCTCGTTCATGACACCATTCCTTGGCGGGCGGTGCTCAACGCGCTCAAC 1828
QY 306 ----- 306
DB 1829 CTACTACTGGGCTGCTCTCTTAATGACAGAGTCCATAGGGAGAGCGTCGACCATGCC 1888
QY 307 -----ThrThrValAla 310
DB 1889 TTGAGAGCCTTCACACCATGACGTCTCGGTGGGGCGGGGACATGACTATGCTGCGCC 1948
QY 311 ThrGlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaVal 330
DB 1949 GCACCTATGACTGCTCTCTTATCATGCAACTGTAAGACAG----- 1990
QY 331 ArgGlnAlaIleThrAlaAlaIle-----LysAla 340
DB 1991 -----GTCCGCGACGCGCTGCGGTCAATTTCCGCGAGACCGCTTGCTGCGAGC 2041
QY 341 AlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAla 360
DB 2042 GCGAGATGATGCGGCTGCTGCGGTATCGGATTCGGATTCGACAGCCCTCGCTCAAGCC 2101
QY 361 IleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsn 380
DB 2102 TTGCTGACGCTGCGCGC-----ACCAACGCTTTC 2131
QY 381 ProLysLeuSerLysValIleSerLeu----- 390
DB 2132 GCGGAGACAGCCATATATCGCGCATGGCGCGCGCGCTGCGCTGCTGCTGCTGCTG 2191
QY 391 -----ThrSerLysTrpValThrValGlyValGlyValValAlaAlaAlaProAla 407
DB 2192 GCGTTCGCGACCGAGCGGATGGCTTCCCATATGATTCCTGCTGCGCGCGC 2251
QY 408 LeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnAsnValAlaGlnPhe 427
DB 2252 ATCGGAGATGCGCGCTGCGCGCATGCTGCGAGCAGTATGATGACGACCATCAGGGA 2311
QY 428 GlnLysGlnValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMet 443
DB 2312 CAGCTTCAA---GGATTCGCTCGCGCTTTACCAAGCTTAATTGATC 2356
RESULT 21
E12544 5658 bp DNA linear PAT 27-APR-1998
LOCUS E12544
DEFINITION Nucleotide sequence of pCPM533 alpha that contains Chlamydia
pneumoniae antigen gene.
ACCESSION E12544
VERSION E12544.1 GI:3251377
KEYWORDS JP 1997015244-A/4.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 5658)
AUTHORS Izutsu,H. and Matsunoto,A.
TITLE MEASURING METHOD FOR ANTICHLAMYDIA-PNEUMONIA ANTIBODY AND REAGENT
THEREFOR AS WELL AS DIAGNOSTIC MEDICINE FOR CHLAMYDIA PNEUMONIA
INFECTION
PATENT: JP 1997015244-A 4 17-JAN-1997;
JOURNAL HITACHI CHEM CO LTD
COMMENT
OS None
OC Artificial sequences.
PN JP 1997015244-A/4
PD 17-JAN-1997
PF 14-MAR-1996 JP 1996057410
PR 28-APR-1995 JP 95P 106011
PI IZUTSU HIROSHI, MATSUMOTO AKIRA

PC G01N33/571,C07K14/295,C07K16/12,G01N33/53,
PC G01N33/569//A61K49/00,C12N15/09;
CC strandedness: Double;
CC topology: Circular;
FH Key Location/Qualifiers
FH FT source 1..5658
FEATURES
source Location/Qualifiers
BASE COUNT 1446 a 1477 c 1428 g 1307 t
ORIGIN
Alignment Scores:
Pred. No.: 1.68e-60 Length: 5658
Score: 1282.00 Matches: 315
Percent Similarity: 62.59% Conservative: 33
Best Local Similarity: 56.65% Mismatches: 75
Query Match: 54.12% Indels: 135
DB: Gaps: 10
US-09-889-314-2 (1-496) x E12544 (1-5658)
QY 1 AspThrAsnMetSerIleSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20
DB 751 GATACAAACATGTCTATTCATCTTCTTCAGGACCTGACAAATCAAAAATATCATGTCT 810
QY 21 GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGln 40
DB 811 CAAGTTCTGACATGCACACCCAGGGCGTCCCAACAAAGATAACTGTCTGCAACGAA 870
QY 41 ThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThr 60
DB 871 ACGAAGCAAAATACACCAACACGTAGGTAAACACATGAGATGAAAGCATGCTCACT 930
QY 61 IleAlaGlyLysSerGlyLysAspLysThrSerSerThrLysThrGlnThrAlaPro 80
DB 931 ATTGCTGCTGCTTGGAAAGACAAACTTCTGACTACAAAACAGAAACACACTCTCA 990
QY 81 GlnGlnGlyValAlaAlaGlyLysGlnSerSerGlnLysAlaGlyAlaAspThr 100
DB 991 CAACAGGAGGTGCTGCTGGGAAAGATCCCAAGTCAAAAGCAAGCGTCTGATCT 1050
QY 101 GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120
DB 1051 GGAGTATCAGGAGCGGCTGCTACTACAGCATCAAAATACGCAAAAATTCCTATGCA 1110
QY 121 ThrSerIleGlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuInsSerLeu 140
DB 1111 ACCCTATTGAAGAGGCGGCAAAAGTATGAGTCTTACCTTAAGTCACTTCAAAAGCCCTC 1170
QY 141 SerAlaAlaGlnMetLysGlnValGlnAlaValAlaValAlaAlaLeuSerGlyLysSer 160
DB 1171 AGTGGCGGCAATGAAAGATCGAAGGCTGTGTGCTGCGCCCTCAGGAAAGT 1230
QY 161 SerGlySerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSer 180
DB 1231 TCGGGTCCGCAAAATTTGAAACACCTGAGCTCCCAAGCCCGGGGTGACACCAAGATCA 1290
QY 181 GlnValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThr 200
DB 1291 GAGGTATATGAAATCGGACTGCGCTGCTTAAGCAATTCAGCATTTGGGAAAGCCACA 1350
QY 201 LysSerAlaLeuSerAsnTyraLaseThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
DB 1351 AATTCGCTTATCTAATCATGACAGTACACAAAGCAAGCAAGCAACAAATTAACATA 1410
QY 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnLysGlnIleMetLys 240
DB 1411 GGTCTAAGAAACAGCGCTAAATAATTCGATTAAGAAGCAAGAAATATACCAAGATGAAG 1470

Pred. No. :	1,68e-60	Length:	5658
Score:	1282.00	Matches:	315
Percent Similarity:	62.59%	Conservative:	33
Best Local Similarity:	56.65%	Mismatches:	75
Query Match:	54.12%	Indels:	135
DB:	6	Gaps:	10

US-09-889-314-2 (1-496) x E12523 (1-5658)

QY	1	ASpThrAsnMeSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSer	20
Db	751	GATCAAAACATGCTATTTCATCTTCCTTCAGACCTGCACATCAAAAAATATCATGTCT	810
QY	21	GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlu	40
Db	811	CAAGTCTGCATCGACACACCCCGAGGGCGTCCCCCAAGAATAAAGCTGCTGCACAACAA	870
QY	41	ThrLysGlnIleGlnGlnThrArgGlnGlnGlyLysAsnThrGlnMetGluSerSpAlaThr	60
Db	871	ACGAAGCAAAATACGCAAAACACGTACAGGATTAATAAACACAGATGGAAAGCCATCCACT	930
QY	61	IleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaPro	80
Db	931	ATTGCTGGTGGCTTTGGAAAAAGACAAATCTCTCCACTACAAAAACAGAAACACTCCA	990
QY	81	GlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThr	100
Db	991	CAACAGGAGTTGCTGCTGGGAAAGAAATCCTCAGAAAGTCAAAAGGACAGTCTGATPACT	1050
QY	101	GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaSerGln	120
Db	1051	GGAGATCAAGAGAGGCGTGTACTACACGATCAAAATACGCAACAAATATTCCTATGACG	1110
QY	121	ThrSerIleGlnGlnAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeu	140
Db	1111	ACCTCTATGAGAGGCGGACCAAAAGATGGAATCTACTCTAGAGTCACTTAAAGCCTC	1170
QY	141	SerAlaAlaGlnMetLysGluValGluAlaValAlaValAlaAlaLeuSerGlyLysSer	160
Db	1171	AGTGCGCGGCAAAATGAAGAAGTCGAAGCGGTGTGTGTCGCCCTCTCAGGAGAAAGT	1230
QY	161	SerGlySerAlaLysLeuGluThrProGluLeuProLysProGlyValThrProArgSer	180
Db	1231	TCGGGTTCCGCAAAATTTGGAAACCTCGAGCTCCCAACGCCGGGGTGACACCAAGATCA	1290
QY	181	GluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThr	200
Db	1291	GAGGTATCGAAATCGGACTCGCGCTTGCTTAAGGCAATTCAGACATTGGGAGAAACCCAA	1350
QY	201	LysSerAlaLeuSerAsnThrYrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu	220
Db	1351	AAATCTGCTTATCTACTACTATGCAAGTACACAGACACAGACAGACCAACCAATTAACATA	1410
QY	221	GlyLeuGlnLysGlnAlaIleLysIleAspLysGluIArgGluGluThrGlnLeuMetLys	240
Db	1411	GGCTCTAGAAAAGCAAGCGAATAAAAATCGATTAAGAAACGAGAAAGAAATACCAAGATGAG	1470
QY	241	AlaAlaGluGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMet	260
Db	1471	GCTGCCACAGAAAGCTTAAGATCTCGAAGGACAAATGATACGTCTCAATCTGTGATG	1530
QY	261	IleAlaValSerValAlaIleThrVal-----IleSer	271
Db	1531	ATMGCGAA- GGGGTTTCGAATTTGCCATGGGGGCCCTTAATTAATTAACTGAGACATCCAG	1589
QY	272	IleValAlaAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeu-----AlaAlaGly	289
Db	1590	ATCTAATCATGATGATCTCTACGCGCGAGCATGTGTGGCGGATCACCGGCCACACAGGT	1649
QY	290	AlaAlaValGlyAlaAla-----	295
Db	1650	GCGGTTGCTGCGGCTTA-TATCGCCGACATCACCGATGGGGAAGATCGGGCTCGCCACTT	1708

OY	344	SerGlyIleLysAlaPheIleLysThrIleValLysAlaIleAlaLysAlaIleSerLys	363
Db	1504	TCTGGAAATPAAAGCATTTATCAAAAACCTTAGTCAAGCGATTGCCAAAGCCATTTCTTAA	1563
OY	364	GlyIleSerLysValPheAlaLysGlyIleThrGlnMetIleAlaLysAsnPheProLysLeu	383
Db	1564	GSAAATCTTAAGGTTTTCGTAAAGGAACTCAAAATGATTCGAAACAACCTCCCAAGCTC	1623
OY	384	SerLysValIleSerSerLeuThrSerLysTrpValIleThrValGlyValValVal	403
Db	1624	TGCAAAAGCATCTCCCTCTTACCGATAATGGGACAGGTGGGGTGGGTTGTAGTT	1683
OY	404	AlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnAsn	423
Db	1684	GGGGGCGCTGCTCTCGGTAAAGGATTTATGCAAAATGCACCTCTCGAGATGCACAATAAC	1743
OY	424	ValAlaGlnPheGlnLysGluValGlyLysLeuGlnAlaIleAlaAspMetIleSerMet	443
Db	1744	GTCGCTCAATTTCAAGAAAGAACTCGGAAACCTGCAGGCTGGCGCTGATATGATTTCTATG	1803
OY	444	PheThrGlnPheTrpGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyIleSerAsn	463
Db	1804	TTCACTCAATTTGGCAACGCAAGATTAATAATTTGCTCAAAAACAACGGCAGATCTAAT	1863
OY	464	GlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTrpAlaIle	483
Db	1864	GAAATGACTCAAAAAGCTACCAAGCTGGGCGCTCAAAATCTTAAAGCGATGCCGCAATC	1923
OY	484	SerGlyAlaIleAlaGlyAla	490
Db	1924	AGCGAGGCATTCGTGGCGCA	1944

RESULT	19
LOCUS	ARI22110
DEFINITION	Sequence from patent US 6165478.
ACCESSION	ARI22110
VERSION	ARI22110.1 GI:14106427
KEYWORDS	.
SOURCE	Unknown.
ORGANISM	Unknown.
REFERENCE	Unclassified. 1 (bases 1 to 5658)
AUTHORS	Izutsu,H., Obara,K. and Matsumoto,A.
TITLE	DNA encoding Chlamydia pneumoniae antigenic polypeptide Patent: US 6165478 A 10-26-Dec-2000;
JOURNAL	Location/Qualifiers
FEATURES	. source 1..5658 /organism="unknown"
BASE COUNT	1446 a 1477 c 1428 g 1307 t
ORIGIN	
Alignment Scores:	
Pred. No.: Score: Percent Similarity: Best Local Similarity: Query Match:	1..68e-60 Length: 5658 1282.00 Matches: 315 62.5% Conservative: 33 56.65% Mismatches: 75 54.12% Indels: 135
DB:	Gaps: 10

Qy	1	AspThrAsnMetSerLleSerSerSerSerGlyProkAspAsnGlnLysAsnLleMetSer	20
Db	751	GATACAAACATGCTATTTCATCTTCTTCAGACCTCAACATCAAAAAATATCATCTCT	81.00
Qy	21	GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGln	40
Db	811	CAAGTCTTGACATGCACACCCCAAGGCGCTGCCCAACAGATAAGCTGTCTGCGACGAA	87.00
Qy	41	ThrLysGlnLleGlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThr	60

Db	871	ACGACGAAATACAGCAACACCTCGGGTATAAAACACTGAGATGGAAGGATGCCACT	930
Qy	61	lIeAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaPro	80
Db	931	ATTGCTGCTCTCTCTGGAAAGACAAACCTCTCGACTACAAAAACGAAACAGCTCCA	990
Qy	81	GlnGlnGlyValAlaAlaGlyLysGlySerSerGlySerGlnLysAlaGlyAlaAspThr	100
Db	991	CAACAGGAGATGCTGCTGGCAAAAGATCTCGAAGACAAAGGCGGTGTCGATACT	1050
Qy	101	GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln	120
Db	1051	GGAGTATCAGAGCGGCTGCTACTACAGCATATAAATCTGCACAAAAATTGCTATGCAG	1110
Qy	121	ThrSerIleGlnGlnAlaSerLysSerMetGlySerThrLeuGlnSerLeuGlnSerLeu	140
Db	1111	ACGCTATTGAAGCGGACCAAAAGTATGAGGTCTACCTTAGATGACTCCAAAGCCTC	1170
Qy	141	SerAlaAlaGlnMetLysGlyValGluAlaValValAlaAlaLeuSerGlyLysSer	160
Db	1171	AGTCCCGCGCAATATGAAGAAAGTCGAAGGCGTGTGTGTGTCGCCCTTCGAGGAAAAGT	1230
Qy	161	SerGlySerAlaLysLeuGluThrProGluLeuProLysProGlyValThrProArgSer	180
Db	1231	TGCGGTTCCGCAAAATGGAAACACTGAGCTCCCAAGCCCGGGGTGACCAAGATCA	1290
Qy	181	GluValIleGluIleGlyLeuAlaLeuAlaAlaLysAlaIleGlnThrLeuGlyGluAlaThr	200
Db	1291	GAGGTATCGAAATCGGACACTCGCGCTTGCTTAAGCAATTCAACACTTGGAGAAAGCCCA	1350
Qy	201	LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu	220
Db	1351	AAACTCGCTTATCTACTACTGCAAGTACCAAGCAAGCAAGCAAGCAAAACAAATAACTA	1410
Qy	221	GlyLeuGluLysGlnAlaIleLysIleAspLysGluArgGlnGluThrGlnGluMetLys	240
Db	1411	GGCTAGAAAGCAAGCGATPAAATATGATPAAAGACGAGAGAAATACCAAGACATGAG	1470
Qy	241	AlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMet	260
Db	1471	GCTCGCGCAAGCAAGTCTAAAGATCTCGAAGAACAAATGATATGTCGAATATGCTGATG	1530
Qy	261	IleAlaValSerValAlaIleThrVal-----IleSer	271
Db	1531	ATCGCGAA- GGGGTTCCGAATTGCCATGGGGGCCCTTAATTATTAATCGAGAGATCCAG	1589
Qy	272	IleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeu-----AlaAlaGly	289
Db	1590	ATCTAATATGATGATCCCTACAGCGCGAGACGATCTGGCCGCGATCACCGGCGCCACAGT	1649
Qy	290	AlaAlaValGlyAlaAla-----	295
Db	1650	GCGGTGCTGGCGCTTA-TATCCGCCACATCACCGATGGGGAAAGATCGGGCTCCGCCTT	1708
Qy	296	-----AlaAlaGlyGlyAla	300
Db	1709	CGGGCTCATGAGCGCTTGTTTTCGGCGTGGGTATGTHGGCAGCGCCGTGGCGGGGACTG	1768
Qy	301	AlaAlaGlyAla-----AlaAlaAla-	306
Db	1769	TTTGGCGCCATCTCTTCATGACGACCATTCCTCTTGGCGGCGGTGCTCAAGCGGCTCAAC	1828
Qy	306	-----	306
Db	1829	CTACTACTGGGCTCTTCTTAATGACAGAGTGCATTAAGGAGACCGCTGCACCATGCGCC	1888
Qy	307	-----ThrThrValAla	310
Db	1889	TTGAGAGCCTTCACACCAAGTCAGTCTCTCCGCTGGGCGCGGGGCGATGACTATGCTGCC	1948
Qy	311	ThrGlnIleThrValGlnAlaValValGlnAlaValLysGlnAlaValIleThrAlaVal	330
Db	1949	GCACCTATGACTGTCTCTTATTCATGCAACTCGTGGAGCAG-----	1990

|||||
Db 1684 GCGGCGCTGCTCGTAAGGATATGCAATGACGCTCGAGATGCAACAAAC 1743
QY 424 ValAlaGlnheGlnySgluValGlyLysLeuGlnAlaAlaAspMetIleSerMet 443
Db 1744 GTGCTCAATTGAGAAAGATCGGAAACTCAGCGCTCGATGATGATTTCATG 1803
QY 444 PheThrGlnPheTrpGlnGlnAlaSerLysIleAlaSerLysGlnPheTrpGlnSerAsn 463
Db 1804 TTCACATCAATTTTGGCAGCAGCAAGTAAATTCCTCCAAACAAACAGCGCACTCAAT 1863
QY 464 GluMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyrAlaIle 483
Db 1864 GAATGACTCAAAAAGCTACCAAGCTGGCGCTCAAAATCCTFAAGCGGTATGCCGCAATC 1923
QY 484 SerGlyAlaIleAlaGlyAla 490
Db 1924 AGCGAGCCATCGCTGCGCA 1944
RESULT 18
E12538
LOCUS E12538 1947 bp DNA linear PAT 27-APR-1998
DEFINITION DNA encoding a fusion protein of DHFR and antigen peptide of
Chlamydia pneumoniae.
ACCESSION E12538
VERSION E12538.1 GI:3251371
KEYWORDS JP 1997015243-A/1.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1947)
AUTHORS Izutsu, H., Obara, K. and Matsumoto, A.
TITLE MEASURING METHOD FOR ANTI-CHLAMYDIA PNEUMONIA ANTIBODY AND REAGENT
THESFOR AS WELL AS DIAGNOSTIC MEDICINE FOR CHLAMYDIA PNEUMONIA
INFECTION
JOURNAL Patent: JP 1997015243-A 1 17-JAN-1997;
HITACHI CHEM CO LTD
COMMENT OS None
OC Artificial sequences.
PN JP 1997015243-A/1
PD 17-JAN-1997
PF 15-MAR-1996 JP 1996058609
PI 28-APR-1995 JP 95P 106009
PR IZUTSU HIROSHI, OBARA KAZUHIKO, MATSUMOTO AKIRA PC
G01N33/569, C12N1/21, C12N15/02, C12P21/02, C12P21/08, PC
G01N33/53
PC G01N33/531, G01N33/571//A61K49/00, C07K14/245, C07K14/39, PC
C07K16/12, C07K19/00,
PC (C12N1/21, C12R1:19), (C12N15/09, C12R1:01), (C12P21/02, C12R1:19);
CC strandedness: Double;
CC topology: Linear;
FH key Location/Qualifiers
FT source 1. 1947
FT mat_peptide 1. 1947
FT /organism='Artificial sequences' FT
FT /product='fusion protein of DHFR and antigen
FT peptide'
FT misc_feature 1. 480
FT /note='DHFR coding region'
FT 484..1947
FT misc_feature /note='antigen peptide coding region'.
FEATURES
source Location/Qualifiers
1. 1947
/organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 591 a 449 c 486 g 421 t
ORIGIN
Alignment Scores: 5.71e-117 Length: 1947
Pred. No.: 2316.00 Matches: 487
Score: 100.00% Conservative: 0
Percent Similarity: 100.00%

Best Local Similarity: 100.00% Mismatches: 0
Query Match: 97.76% Indels: 0
DB: 6 Gaps: 0
US-09-889-314-2 (1-496) x E12538 (1-1947)
QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValIleu 23
Db 484 ATGCTATTTCATCTTCTTGAGGACCTGACATCAACAAAAATATCATGTCTCAAGTTCTG 543
QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln 43
Db 544 ACATCGACACCCAGGCGGCGCCCAAGATAGCTGCTGCGCAAGCAAGCAAGCA 603
QY 44 IleGlnGlnThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThrIleAlaGly 63
Db 604 ATACAGCAAAACAGCTCAGGCTGAAACACAGATAGCAAGCAAGCAAGCAAGCAAGCA 663
QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83
Db 664 GCTTCTGGAAGCAACAACTCTCTGACACTACAAAAACGAAACAGCTCAGCAACAGCA 723
QY 84 ValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyValSer 103
Db 724 GTTGCTGCTGGAAAGAAATCCTCAGAAAGTCAAAAGCAGCTGCTGATCTGAGATATCA 783
QY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
Db 784 GGAGCGCTGCTACTACAGATCAAAATACGCAACAAATTCCTATGCAAGCTCTATAT 843
QY 124 GlnGlnLysSerLysSerMetGluSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143
Db 844 GAAGAGCGGAGCAAAAGTACGAGCTACTGAGTACTGAGTACTGAGTACTGAGTACTGAG 903
QY 144 GlnMetLysGluValGlnAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlySer 163
Db 904 CAATGAAAGAAAGTCAAGGAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 963
QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGluValIle 183
Db 964 GCAAAATGGAACACCTGAGCTCCCAACCCGGGGTGACACCAAGATCAGAGGTATAC 1023
QY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyLysAlaThrLysSerAla 203
Db 1024 GAATGCGAGCTCGCGCTGCTGTAAGCATTCACACATTTGGGGAACACCAAAATCTGCC 1083
QY 204 LeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlu 223
Db 1084 TTATCTAACTATGCAAGTACACAGCAACAGCACCAACCAAAATTAAGTCTAGAA 1143
QY 224 LysGlnAlaIleLysIleAspLysGluArgGlnGlnGlnGlnGlnGlnGlnGlnGln 243
Db 1144 AACCAAGCGATTAACATCGATTAAGAAAGCAAGAAATACCAAGATGAAAGCTGCCGAA 1203
QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaVal 263
Db 1204 CAAAGCTTAATATCTCGAAGACATGATACGTCTCATCTGATGATGATGATGATGAT 1263
QY 264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu 283
Db 1264 TCTGTGCTCATTCAGATTATTTCTATTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1323
QY 284 AlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaAlaAlaAlaAlaAlaAla 303
Db 1324 GCTGAGCTGCTCGGAGCTGCTGAGTACGAGCGGACCTGAGAGTGCAGCAGAGAGCT 1383
QY 304 AlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLys 323
Db 1384 GCTGCGCAACCAAGCTAGCAGCAACAATTCAGCTTCAGAGGCTGAGGAGGAGGAG 1443
QY 324 GlnAlaValIleThrAlaValArgAlaAlaIleThrAlaAlaIleLysAlaAlaValLys 343
Db 1444 CAAGCTGTATTCACAGCTGTCAGCAACAGCATACCGCGGCTATTAAGAGGCTGTGCANA 1503

DB 1924 AGCGAGCCATCGTCGGCA 1944

RESULT 17

LOCUS E12525

DEFINITION DNA encoding a fusion protein of DHFR and antigen peptide of Chlamydia pneumoniae.

ACCESSION E12525

VERSION E12525.1 GI:3251358

KEYWORDS JP 1997009976-A/L.

SOURCE unclassified.

ORGANISM unclassified.

REFERENCE 1 (bases 1 to 1947)

AUTHORS Izutsu, H., Obara, K. and Matsumoto, A.

TITLE DIHYDROFOLIC ACID REDUCTASE-CHLAMYDIA PNEUMONAE ANTIGEN FUSED PROTEIN, DNA CODING FOR THE SAME, RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT CONTAINING THE VECTOR, AND PRODUCTION OF ANTI-CHLAMYDIA PNEUMONAE ANTIBODY

JOURNAL Patent: JP 1997009976-A 1 14-JAN-1997;

HITACHI CHEM CO LTD

COMMENT OS None

OC Artificial sequences.

PN JP 1997009976-A/1

PD 14-JAN-1997

PF 14-MAR-1996 JP 1996057409

PR 28-APR-1995 JP 95P 106006

PI IZUTSU HIROSHI, OBARA KAZUHIKO, MATSUMOTO AKIRA PC

C12N15/09, C07H21/04, C07K14/295, C07K14/47, C07K19/00, C12N1/21, PC

C12P21/02, PC

GO1N33/569, GO1N33/571//A61K39/395, A61K49/00, C12P21/08, PC

(C12N15/09, C12R1:01), PC

(C12N1/21, C12R1:19), (C12P21/02, C12R1:19), (C12P21/02, C12R1:91), PC

(C12P21/08, PC

C12R1:91);

CC Strandedness: Double;

CC topology: Linear;

FH Key Location/Qualifiers

FT source 1. 1947

mat_peptide 1. 1947

FT /organism='Artificial sequences' FT

FT /product='fusion protein of DHFR and antigen peptide'

FT misc_feature 1. 480

FT /note='DHFR coding region'

FT misc_feature 484..1947

FT /note='antigen peptide coding region'.

FEATURES

source 1. 1947

Location/Qualifiers

/organism='unclassified'

/db_xref='taxon:32644'

BASE COUNT 591 a 449 c 486 g 421 t

ORIGIN

Alignment Scores:

Pred. No.: 5 71e-117 Length: 1947

Score: 2316.00 Matches: 487

Percent Similarity: 100.00% Conservative: 0

Best Local Similarity: 100.00% Mismatches: 0

Query Match: 97.76% Indels: 0

DB: 6 Gaps: 0

US-09-889-314-2 (1-496) x E12525 (1-1947)

QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnIleAsnIleMetSerGlnValLeu 23

DB 484 ATGCTGATTTCACTCTCTCAGAGACTGACAAATCAAAAAATATCATGCTCAAGTTTCG 543

QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyValAsnGluThrLysGln 43

DB 544 ACATGCACACCCAGGCGTGCCTCCCAACAAGATTAAGCTGTCTGGCAACGAAGCAAGCA 603

QY 44 ILeGlnGlnThrArgGlnGlyLysAsnThrGlnMetGluSerAspAlaThrIleAlaGly 63

DB 604 ATACAGCAACACCGTCAGGCTAAACACTGAGATGAGAAAGCGATGCCATATTGCTGCT 663

QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGly 83

DB 664 GCTTCTGGAAGAAACAACTTCTCGACTACAAAAGCAAGAAACAGCTCCACACAGGGA 723

QY 84 ValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyValSer 103

DB 724 GTTGCTGTGGGAAAGAACTCCACAGAAAGTCAAAAGCGGTGTGATCTGAGATATCA 783

QY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123

DB 784 GGAGCGGCTGCTACTACACATCAATTAAGTCAACAAATTTGCTATGACACCTTATT 843

QY 124 GluGluAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeuSerAlaAla 143

DB 844 GAAAGGCGCAGCAAAAGTATGAGTCTACCTTAGAGTCACTTCAAAAGCCTCAGTGGCGG 903

QY 144 GlnMetLysGluValGlnAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlySer 163

DB 904 CAAATGAAAGAAAGTCAAGCGGTGTGTGCTGCTCCCTCAGGGGAAAGTTTCGGGTTCC 963

QY 164 AlaLysLeuGluThrProGluLeuProLysProGlyValThrProArgSerGluValIle 183

DB 964 GCAAATTTGGAACACCTGAGTCTCCCAAGCCGGGGGTACACCAAGATCAAGAGTTATC 1023

QY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThrLysSerAla 203

DB 1024 GAAATCGGACCTGCGGCTCTTAAGCAATTCAGACATTTGGAGAAAGCCAAATTCGCC 1083

QY 204 LeuSerAsnThrLysThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlu 223

DB 1084 TTATCTTAATTAATGACATGACCAAGCAGCAGCAAGCAAACTAATAGCTGTAGAA 1143

QY 224 LysGlnAlaIleLeuSylLeuAspLysGluArgGluGluTyrGluGluMetLysAlaAlaGlu 243

DB 1144 AGCAAGCGATTAATAATGATTAAGACAGAGAGATACCAAGATGAAGCTGCGGAA 1203

QY 244 GlnLysSerLysAspLeuGluGlyThrMetAspThrValAsnThrValMetIleAlaVal 263

DB 1204 CAGAAGCTTAAGATCTCAAGAGCAACAAATGATGATCAATGATGATGATGATGATG 1263

QY 264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu 283

DB 1264 TCTGTTCATTAACAGTATTTCTATTGTTGCTGCTATTATTACATGCGGACTGACTC 1323

QY 284 AlaGlyLeuAlaAlaGlyAlaValAlaGlyAlaAlaAlaAlaGlyAlaAlaGlyAla 303

DB 1324 GCTGACTGCTGCGGAGACTGCTGAGTGCAGCGGCGAGCTGGAGGTGCACAGAGACT 1383

QY 304 AlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLys 323

DB 1384 GCTGCGGCAACCGGTACCAACAAATTAAGTCAAGCTGTGTGTCCAAACGGGAA 1443

QY 324 GlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaValLys 343

DB 1444 CAAAGCTTATACAGCTCTCAGCAAGCAAGTACCGCGCTATTAACCGCTGTCAAA 1503

QY 344 SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363

DB 1504 TCTGGAATAAAGCATTTTTCAAAACCTTAGTCAAAAGCATTTGCCAAAGCCATTTCCTAAA 1563

QY 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383

DB 1564 GGAATCTTAAGGTTTTCCTTAAGGAACTCAATGATTTGCCAAGAACTTCCCAAGCTC 1623

QY 384 SerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValGlyValVal 403

DB 1624 TCGAAAGTATCTGCTCTTACAGTAATAGGTCACGTTGGGGTGTGGTGTGTGATTT 1683

QY 404 AlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGluMetGlnGlnAsn 423

Db 1201 GCGGCGCTCTCTCGGTAAAGGATTATGCAATTCAGCTCTCGAGATGCACAAAC 1260
QY 424 ValAlaInPheGlnYsGluValGlyLysLeuGlnAlaAlaAspMetIleSerMet 443
Db 1261 GTGCTCAATTTGAGAAAGATCGAAACCTGCAGCTGCGCTGATATGATTCTATG 1320
QY 444 PheHrGlnPheThrPProGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyLysAsn 463
Db 1321 TTGACTCAATTTTGCAACAGGCAAGTAAATTCCTCAAAACAAACAGGCACTTAAT 1380
QY 464 GluMetHrGlnYsAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyrAlaAlaIle 483
Db 1381 GAAATGACTCAAAAGCTACCAACGCTGGCGCTCAAAATCCTTAAGCGTATGCCGCAATC 1440
QY 484 SerGlyAlaIleAlaGlyAla 490
Db 1441 AGCGAGCCATCGCTGCGCA 1461
RESULT 16
AR122114
LOCUS AR122114 1947 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 17 from patent US 6165478.
ACCESSION AR122114
VERSION AR122114.1 GI:14106431
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1947)
AUTHORS Izutsu,H., Obara,K. and Matsumoto,A.
TITLE DNA encoding Chlamydia pneumoniae antigenic polypeptide
JOURNAL Patent: US 6165478-A 17 26-DEC-2000;
FEATURES
source location/Qualifiers
1..1947
BASE COUNT 591 a 449 c 486 g 421 t
ORIGIN
Alignment Scores:
Pred. No.: 5,71e-117 length: 1947
Score: 2316.00 Matches: 487
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 97.76% Indels: 0
DB: 6 Gaps: 0
US-09-889-314-2 (1-496) x AR122114 (1-1947)
QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23
Db 484 ATGCTATTTCATCTTCTTCAGACCTGACAAATCAAAAAATATCATGTCTCAAGTTCTG 543
QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln 43
Db 544 ACATCGACACCCAGGGGCTGCCCAACAAATAGCTGTGGCAACGAACGAACGA 603
QY 44 IleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63
Db 604 ATTAACGAACACGTCAGGATTAACACTGACATGCAAAACGATGCCACATTTGCTGTG 663
QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83
Db 664 GCTTCTGGAAGAACAAACTTCTCGACTACAAAACAGAAACAGACGTCACCAACAGGGA 723
QY 84 ValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyValSer 103
Db 724 GTTCTGCTGGGAAAGATCTCTCAGAAAGTCAAAAGGAGGCTGCTGATCTGAGTATCA 783
QY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
Db 784 GGAGCGGCTGCTACTACAGCATCAAAATCTGCAACAAAAATTTGCTATGACAGCTCTATT 843
QY 124 GluGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerAlaAla 143

Db 844 GAAAGAGCGAGCAAAAGTATGAGACTTACCTTAGAGTCACCTCAAAAGCTCAGTGCAGCG 903
QY 144 GlnMetLysGluValAlaGluAlaValAlaValAlaLeuSerGlyLysSerSerGlySer 163
Db 904 CAAATGAAAGAAAGTCGAAGGGGTGTGTCGCCCTCTCAGGAAAGTATGGGGTTC 963
QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGluValIle 183
Db 964 GCAAAATTTGCAACACCTGAGCTCCCAAGCCCGGGGTGACACCAAAATCAGAGATTATC 1203
QY 184 GluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlyLysAlaThrLysSerAla 203
Db 1024 GAAATCGAGCTCGCGCTTGCTTAAGCAATTCAGACATTTGGAGAGCAACAAATATCGCC 1083
QY 204 LeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlu 223
Db 1084 TTATCTAATCTATGCAATGACAAAGCACAGCACACCAACAAATTAACATGATCTAGAA 1143
QY 224 LysGlnAlaIleLysIleAspLysGluArgGluGlnThrGlnGlnMetLysAlaAlaGlu 243
Db 1144 AAGCAGCGATTAATAATCGATTAAGAAACGACAGAAATACCAAGATGAAGGCTGCCGAA 1203
QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaVal 263
Db 1204 CAGAAAGTCTTAAGATCTCGAAGGAACAATGATGATCTGCAATACTGTGATGATCGCGTT 1263
QY 264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu 283
Db 1264 TCTGTTCGCATTAACAGTATTTCTATTGTGCTCTCTATTATTCATCGGAGCGTCACTC 1323
QY 284 AlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaAlaAlaGlyAlaAlaGlyAla 303
Db 1324 GCTGAGCTCGCTCGCGGAGCTGCTGTAGGTGCACGCGCAGCTGAGGTGCACGAGAGCT 1383
QY 304 AlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLys 323
Db 1384 GCTGCGCAACCAACGAGTACCAACAAATATACATTCAGCTGTTGTCACCAACGCGTGAA 1443
QY 324 GlnAlaValAlaIleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLys 343
Db 1444 CAACCTGTTACACAGCTGTGACCAAGCAAGCAACGCGCGCTATTAAGCGCGTCAAA 1503
QY 344 SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363
Db 1504 TCTGAATTAAGCATTTATTAACAACTTACTCAAGCGATTGCCAAGCATTTCTTAA 1563
QY 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383
Db 1564 GGAATCTCTAAGTTTTCGCTAAGGAACTCAAAATGATTGCGAAGAACTTCCCAAGCTC 1623
QY 384 SerLysValIleSerSerLeuThrSerLysTyrValThrValGlyValGlyValValVal 403
Db 1624 TCGAAAGTCATCTGCTCTTACCAAGTAATGGGTCAAGGGTGGGGTGTGTGTGTGT 1683
QY 404 AlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnGlnAsn 423
Db 1684 GCGGCGCTGCTCGTGGTAAGGATTAATGCAAAATGCAAGCTCTCGAGATGCACCAAAAC 1743
QY 424 ValAlaGlnPheGlnLysGluValAlaGlyLysLeuGlnAlaAlaAlaAspMetIleSerMet 443
Db 1744 GTCCCTCAATTTGAGAAAGATGCGAAMAACCTGCAGGCTGCGCTGATATGATTCTATG 1803
QY 444 PheHrGlnPheThrPProGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyLysSerAsn 463
Db 1804 TTCACTCAATTTTGGCAACAGCAAGTAAATTCCTCAAAACAAACAGGCACTCTTAAT 1863
QY 464 GluMetHrGlnYsAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyrAlaAlaIle 483
Db 1864 GAAATGACTCAAAAGCTACCAACGCTGGCGCTCAAAATCCTTAAGCGTATGCCGCAATC 1923
QY 484 SerGlyAlaIleAlaGlyAla 490
|||||

Db 1321 TTCACGCAATTTGGCAACAGGCAAGTAAATTCCTCAAAACAAACAGCGACTAAT 1380
QY 464 GUmethThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaThrAlaIle 483
Db 1381 GAAATGACGCAAAAGCTACCAAGCTGGCGCTCAATCTTAAAGCGTATGCCGAATC 1440
QY 484 SerGlyAlaIleAlaGlyAla 490
Db 1441 AGCGAGCGCATCGTCGGCCA 1461
RESULT 15
E16635 1464 bp DNA linear PAR 28-JUL-1999
LOCUS Chlamydia pneumoniae gene for 53kDa-antigen peptide.
DEFINITION E16635
ACCESSION E16635
VERSION E16635.1 GI:5711318
KEYWORDS JP 198210978-A/1.
SOURCE unclassified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1464)
AUTHORS Izutsu,H. and Hagiwara,T.
TITLE RECOMBINANT VECTOR AND TRANSFORMANT CONTAINING THE SAME, AND
RECOMBINANT VACUOLAR VIRUS AND ITS PRODUCTION, AND PRODUCTION OF
CHLAMYDIA PNEUMONIAE ANTIGEN POLYPEPTIDE
PATENT: JP 198210978-A 1 11-AUG-1998:
HITACHI CHEM CO LTD
COMMENT OS Chlamydia pneumoniae
PN JP 198210978-A/1
PD 11-AUG-1998
PF 31-JAN-1997 JP 1997018523
PI IZUTSU HITROSHI, HAGIMARA TOSHIKATSU
PC C12N15/09,C07H21/04,C07K14/235,C12N1/21,C12N5/10,C12N7/00, PC
C12P21/02//
PC A61K39/00,A61K39/118,C12Q1/68,(C12N15/09,C12R1:01),(C12N1/21,
PC C12R1:19),
PC (C12N5/10,C12R1:91),(C12N7/00,C12R1:92),(C12P21/02,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
FH key Location/Qualifiers
FT source 1..1464
FT /strain='YK-41' /organism='Chlamydia pneumoniae' FT
FT mat_peptide 1..1464 /tissue_type='elementary bodies' FT
FT Location/Qualifiers
1..1464
/db_xref='taxon:32644'
BASE COUNT 470 a 332 c 348 g 314 t
ORIGIN
Alignment Scores:
Pred. No.: 4.05e-117 Length: 1464
Score: 2316.00 Matches: 487
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 97.76% Indels: 0
DB: 6 Gaps: 0
US-09-889-314-2 (1-496) x E16635 (1-1464)
QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23
Db 1 ATGCTATATTCACCTCTCTCAGAGCACTGACAAATCAAAAAATATCATGCTCAAGTTCTG 60
QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerSerGlyAsnGlnThrLysGln 43
Db 61 ACATTCACACCCCGAGGCGCGCCCAACAAAGATTAAGCTGTCTGGCAACGAAACGAAAGCA 120
QY 44 IleGlnIleThrArgGlnGlyLysAsnThrGlnMetGlnUserSerAspIleThrIleAlaGly 63

Db 121 ATACAGCAAAACGCTCAGGCTAAAAAACACTGAGATGGAAGGAGGATCCATCTGTGCT 180
QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83
Db 181 GCTTCTGGAAAAAGACAAAACCTCTCGACTACAAAACGAAACAGCTCCACACAGGCA 240
QY 84 ValAlaAlaGlyLysGlnUserSerGlnUserGlnLysAlaGlyAlaAspThrGlyValSer 103
Db 241 GTTGCTGCTGGGAAGAAATCTCAGAAAGTCAAAAGCGAGGTGCTGATCTGAGATATCA 300
QY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
Db 301 GGAGCGCGCTGCTACTACACATCAAAATAGTGCACAAAAATGCTATGACAGACTCTATT 360
QY 124 GlnGlnAlaSerLysSerSerMetGlnUserThrLeuGlnUserGlnSerLeuSerAlaIle 143
Db 361 GAAGAGCGCAGCAAAAGATGATGAGTCTACCTTAAGATCACTTCAAAAGCTTCAGTCCGCG 420
QY 144 GlnMetLysGlnValGlnAlaValAlaValAlaIleUserSerGlyLysSerSerGlySer 163
Db 421 CAAATGAAGAAGTCGAACGGGTGTGTGCTGCTCAGGGAAAGTTCCGGGCTCC 480
QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnValIle 183
Db 481 GCAAAATTTGAAACACCTGAGCTCCCAAGCCCGGGGTGACACCAAGATCAGAGTTATC 540
QY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThrLysSerAla 203
Db 541 GAAATCGGACCTCGCGCTCTCTTAAGCAATTCAGACATTGGGAGAAAGCCCAAAATCTGCC 600
QY 204 LeuSerAsnThrLysSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnGly 223
Db 601 TTTATCTACTATGCAAGTACACAGCAAGCAAGCAACCAAAATTAAGTCTAGTCTAGAA 660
QY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 243
Db 661 AAGCAAGCATTAATAATCATTAAGAAAGAAAGAAAGATATCAAGATGAAGCTGCCGAA 720
QY 244 GlnLysSerLysAspLeuGlnGlnThrMetAspThrValAsnThrValMetIleAlaVal 263
Db 721 CAGAACTCTTAAGATCTCCAGAGCAAAAGATGATGATGATGATGATGATGATGATGAT 780
QY 264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlnAlaGlyLeu 283
Db 781 TCTGTTCCTATTCACATTAATTTCTTATTTGCTGCTATTTTTCATTCGCGAGCTGACTC 840
QY 284 AlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaAlaAlaAlaGlyGlyAlaAlaGlyAla 303
Db 841 GCGTGACTCGTGGCGGAGCTGCTGATGAGTGCAAGCGGAGCTGAGAGTCCAGAGAGCT 900
QY 304 AlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLys 323
Db 901 GCTGCCGCAACCGCTGACACACAAATTAAGTCAAGCTGTGCTCCAAAGCGGGGAAA 960
QY 324 GlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLys 343
Db 961 CAAAGCTGTTATCACAGCTGCGACAGCAAGATACCGCGCTATTAAGCGGCTGCAAA 1020
QY 344 SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363
Db 1021 TCTGGAATAAAGCATTTATCAAAACTTTAGTCAAAAGCGATTCGCAAAACCATTTCTAAA 1080
QY 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383
Db 1081 GGAATCTTAAGGTTTTCCTTAAGGCAACTCAAAATATTCGGAAGAACTTCCCAAGCTC 1140
QY 384 SerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValGlyValVal 403
Db 1141 TCGAAAGTATCTCGTCTCTTACAGTAAATGGGTACGGTTGGGGTTGGAGTTTGAAGTT 1200
QY 404 AlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnGlnAsn 423

QY 484 SerGlyAlaIleAlaGlyAla 490
Db 1441 AGCGAGCCATCGCGGCCA 1461
RESULT 14
LOCUS E12541 1464 bp DNA linear PAT 27-APR-1998
DEFINITION DNA encoding an antigen peptide Chlamydia pneumoniae.
ACCESSION E12541
VERSION E12541.1 GI:3251374
KEYWORDS JP 1997015244-A/1.
SOURCE JP 1997015244-A/1.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1464)
AUTHORS Iutsu, H. and Matsmoto, A.
TITLE MEASURING METHOD FOR ANTICHLAMYDIA-PNEUMONIA ANTIBODY AND REAGENT THEREFOR AS WELL AS DIAGNOSTIC MEDICINE FOR CHLAMYDIA PNEUMONIA INFECTION
JOURNAL Patent: JP 1997015244-A 1 17-JAN-1997;
COMMENT HITACHI CHEM CO LTD
OS Chlamydia pneumoniae
PN JP 1997015244-A/1
PD 17-JAN-1997
PF 14-MAR-1996 JP 1996057410
PR 28-APR-1995 JP 95P 106011
PI IZUTSU HIROSHI, MATSUMOTO AKIRA
PC GOIN33/571, COTK14/295, COTK16/12, GOIN33/53,
CC GOIN33/569//M61K49/00, C12N15/09;
strandedness: Double;
CC topology: linear;
FH Key Location/Qualifiers
FT source 1. 1464
FT FT
FT /strain='YK41', /organism='Chlamydia pneumoniae' FT
FT mat_peptide 1. 1464
FT /product='antigen peptide',
FEATURES
source Location/Qualifiers
1. 1464
/organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 470 a 331 c 348 g 315 t
ORIGIN
Alignment Scores:
Pred. No.: 4.05e-117 Length: 1464
Score: 2316.00 Matches: 487
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 97.76% Indels: 0
Gaps: 0
US-09-889-314-2 (1-496) x E12541 (1-1464)
QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValIleu 23
Db 1 ATGTCTATTTCATCTCTTCAGACCTGACAAATCAAAAAATTCATGCTCAAGTTCTG 60
QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln 43
Db 61 ACATTCACACCCGAGCGCGGCCCAACAAGATAGCTGCTGCGCAACGAAGCAAGCA 120
QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnUserAspAlaThrIleAlaGly 63
Db 121 ATTCACCAACACGTCAGGTTAAACACCTGACATGAGCAAGCAAGCATGCCATTGCTGGT 180
QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGln 83
Db 181 GCTTCGGAAGAAACAAACTTCTCGATACCAAAAACGAACACCTCCACACACAGGA 240
QY 84 ValAlaAlaGlyLysGlnSerSerSerGlnLysAlaGlyAlaAspThrGlyValSer 103

Db 241 GTTCTGCTGGGAAGATCTTCAGAAAGTCAGAAAGCGAGTGTGATCTGAGATATCA 300
QY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
Db 301 GGAGCGCTGCTACTACAGCATCAATATCTGCAACAAAATTCGTATGACAGACTTATT 360
QY 124 GlnGlnLysSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143
Db 361 GAAGAGCGAGCAAAAGATGAGTACCTTACAGTCACTTCAAAAGCTCAGGCGCG 420
QY 144 GlnMetLysGlnValGlnAlaValAlaAlaLeuSerGlyLysSerSerGlySer 163
Db 421 CAATGAAAGACGCGAAGCGGTTGTTGCTGCTGCTTCAGGAAAAGTTCGGGTTCC 480
QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlyValIle 183
Db 481 GCAAAATTTGAAACACCTGAGCTCCCAAGCCCGGGGTGACACCAAGATCAGAGTTATC 540
QY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThrLysSerAla 203
Db 541 GAAATCGAGCTCGGCTTGTAAAGCAATTCAGACATTCGGAAGCCACAAAATCTGCC 600
QY 204 LeuSerAsnTyraLysSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGln 223
Db 601 TTATCTACTATGCAATGATACACAGCAAGCAGACCAACAAATTAATAGTCTTAGA 660
QY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnLysGlnLysAlaAlaGln 243
Db 661 AAGCAACGATTAATAATTCGATTAAGAAAGAGAAATACCAAGATGAAGCTGCGCAA 720
QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaVal 263
Db 721 CAGAACTTAAGATTCGAGGAACATGATATCTCATCTGATGATCGCGGCTT 780
QY 264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu 283
Db 781 TCTGTTGCCATTACAGTATTCTATTGCTGCTGCTATTTTTTCATGCGGAGCTGGACTC 840
QY 284 AlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAlaAlaAlaGlyAlaAlaGlyAla 303
Db 841 GCTGAGCTCGCTCGGAGCTGCTGAGTGCAGCGGCGAGCTGAGAGTGCAGAGAGCT 900
QY 304 AlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValGlnAlaValLys 323
Db 901 GCTGCCGCAACCGAGTAGCAACAAATTAAGTTCAAGTCTGTTCCAAAGCGGTGAAA 960
QY 324 GlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLys 343
Db 961 CAAAGCTTATACACACTGTCACACACAGCATCACCGCGGCTTTAAAGCGGCTGTCAAA 1020
QY 344 SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363
Db 1021 TCTGCAATTAAGCATTTATCAAAACTTATGTCAAAGCATTCGCCAAAGCAATTTCTTAA 1080
QY 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383
Db 1081 GGATCTCTAAGGTTTTCCTAAGGGAACCTCAAAATGATTCGCAAGAACTCCCAAGCTC 1140
QY 384 SerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValAlaValAla 403
Db 1141 TCGAAAGTATCTCGCTCTTACCACTAAATGGGTACCGGTTGGGTTGTAGTT 1200
QY 404 AlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnGlnAsn 423
Db 1201 GCGGCGCTGCTCTCGTTAAAGGATTAATGCAATGACACTCTCGGAGATGCAACAAAAC 1260
QY 424 ValAlaGlnPheGlnLysGlnValAlaGlyLysLeuGlnAlaAlaAspMetIleSerMet 443
Db 1261 GTGCTCAATTTTCAGAAAGATCGGAAACATGCAAGGCTCGGATATGATTTCTATG 1320
QY 444 PheThrGlnPheTrpGlnAlaSerLysIleAlaSerLysGlnThrGlyLysSerAsn 463

LOCUS E12535 1464 bp DNA linear PAT 27-APR-1998
 DEFINITION Chlamydia pneumoniae 53kD-antigen gene.
 ACCESSION E12535
 VERSION E12535.1 GI:3251368
 KEYWORDS JP 1997009999-A/1.
 SOURCE unclassified.
 ORGANISM unclassified.
 REFERENCE 1 (bases 1 to 1464)
 AUTHORS Izutsu, H. and Matsumoto, A.
 TITLE PROBE AND PRIMER FOR MEASURING CHLAMYDIA PNEUMONIAE GENE.
 JOURNAL MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE USING THE SAME PROBE OR
 PRIMER AND REAGENT FOR MEASURING CHLAMYDIA PNEUMONIAE GENE
 Patent: JP 1997009999-A 1 14-JAN-1997;
 HITACHI CHEM CO LTD
 COMMENT OS Chlamydia pneumoniae
 PN JP 1997009999-A/1
 PD 14-JAN-1997
 PE 15-MAR-1996 JP 1996058608
 PR 28-APR-1995 JP 95P 106008
 PI IZUTSU HIROSHI, MATSUMOTO AKIRA
 PC C1201/68, C07H21/04, C12N1/21, C12N15/02, C12N15/09, C12P21/08, PC
 G01N33/366,
 PC G01N33/569, G01N33/571//A61K49/00, C07K14/295, C1201/04,
 (C12N1/21, PC C12R1:19);
 CC strandedness: Double;
 CC topology: Linear;
 FH key Location/Qualifiers
 FT source 1. 1464
 FT /strain='YK-41',
 FT mat_peptide 1. 1464
 FT /product='53kD-antigen peptide',
 FT Location/Qualifiers
 1. 1464
 /organism='unclassified'
 /db_xref='taxon:32644'
 BASE COUNT 470 a 331 c 348 g 315 t
 ORIGIN
 Alignment Scores:
 Pred. No.: 4.05e-117 Length: 1464
 Score: 2316.00 Matches: 487
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 97.76% Indels: 0
 DB: Gaps: 0
 US-09-889-314-2 (1-496) x E12535 (1-1464)
 QY 4 MetSerLleSerSerSerSerGlyProAspAsnGlnLysAsnLleMetSerGlnValLeu 23
 DB 1 AAGTTATATTTTCATCTTCTTCAGAGACCTGCACATCAAAAAATATCATGTCTCAAGTTCTG 60
 QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGln 43
 DB 61 ACATGACACACCCAGGCGCTGCCCAACAAGTAAAGCTTCTGGCAGCAAGCAAGCAA 120
 QY 44 IleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63
 DB 121 ATACAGCAAAACGTCAGGCTAAACAACTGAGATGAAAGGATGCCACTTGTGCTGT 180
 QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83
 DB 181 GCTTGTGAAAAGACAACTTCCTGACTACAAAAAGCAAGAGCTCCACACAGGGA 240
 QY 84 ValAlaIaGlnPheGlnLysGlnSerSerGlnLysAlaGlyAlaAspThrGlyValSer 103
 DB 241 GTTGGCTGGGGAAGAAATCCTCAGAAAGTCAAAAGCGAGGTGCTGATCTGAGATATA 300
 QY 104 GlyAlaAlaIaIaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
 |||||||

DB 301 GGAGCGGCTGCTACTACAGCATCAAAATACGCCAACAAAAATTTGCTATGCAGACCTCTATT 360
 QY 124 GlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerLysAla 143
 DB 361 GAAGAGCGGAGCAAAAGTAGTAGTCTACCTTAGATCACTTCAAGACCTCAGTCCGGG 420
 QY 144 GlnMetLysGlnValAlaGlnAlaValAlaValAlaAlaLeuSerGlyLysSerSerLys 163
 DB 421 CAAATGAAAGAGTCGAGAGCGGTTGTGTGCTGCCCTCTCAGGAAAAAGTTGGGTTCC 480
 QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnValIle 183
 DB 481 GCAAAATTTGAAACACCTGAGCTCCCAAGCCCGGGGTACACCAAGATCAAGAGTTTATC 540
 QY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlnAlaThrLysSerAla 203
 DB 541 GAAATCGGACCTGCGCTTGTCTTAAGCAATTCAGACATTTGGGAAAGCCCAAAATCTGCG 600
 QY 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnLys 223
 DB 601 TTATCTACTATGCAAGTACACAGCAGACAGCAAGCAACCAAAATTAAGTACTAGTCTAGA 660
 QY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGln 243
 DB 661 AAGCAAGCGATAAATGATTAAGACAGAGAAAGATACCAAGATGAAAGCTGCCGCA 720
 QY 244 GlnLysSerLysAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 263
 DB 721 CAGAAGTCTAAAGATCTCAAGAGCAAGATGACTGCAATATGTAATGATCCGGTT 780
 QY 264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu 283
 DB 781 TCTGTGTCATTCAGTATTTCTATTGTGCTGCTATTTTACATGCGAGCTGAGAC 840
 QY 284 AlaGlyLeuAlaIaGlyAlaAlaValAlaGlyAlaAlaAlaAlaGlyAlaAlaGlyAla 303
 DB 841 GCTGACCTGCTGCGGAGAGCTCTAGCTGACGCGGAGCTGAGGTGACAGAGAGCT 900
 QY 304 AlaAlaAlaIaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaLys 323
 DB 901 GCTGCGGCAACCGATGCAACACCAAAATTCAGTTCAGTGTGTGTCACAGCGGGA 960
 QY 324 GlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaAlaLys 343
 DB 961 CAAGCTGTATATACAGCTGTACAGACAGATCACCGCGCTATTAAGCGGCTGCAAA 1020
 QY 344 SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363
 DB 1021 TCTGAAATFAAAGCATTTTATCAAAACTTTCAGCAAAAGCATTCGCAAAATCTTAA 1080
 QY 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383
 DB 1081 GGAATCTCTTAAGTTTTCCTTAAGGAACTCAAAATGATTGCAAGAACTTCCCAAGCTC 1140
 QY 384 SerLysValIleSerSerLeuThrSerLysTrpValIleThrValGlyValAlaVal 403
 DB 1141 TCGAAAGTATCTCGTCTTACCACTAAATGGTACAGGTTGGGTTGGGTTGTAGTT 1200
 QY 404 AlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnGln 423
 DB 1201 GCGGCGCTGCTCTCGGTAAAGGATTATGCAAAATGACACTCTCGAGATGCAACAAAAC 1260
 QY 424 ValAlaGlnPheGlnLysGlnValAlaGlyLysLeuGlnAlaAlaAlaAspMetIleSerMet 443
 DB 1261 GTGCTCAATTTTCAGAAACATCGGAAACATGCAAGGCTGCGGTGATATGATTTCTAAG 1320
 QY 444 PheThrGlnPheThrProGlnAlaSerLysIleAlaSerLysGlnThrGlyLysLeuSer 463
 DB 1321 TTCACGCAATTTTGGCAACAGCAAGTAAATTTGCTCAAAACCAACAGCGAGCTTAAT 1380
 QY 464 GlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaThrAlaIle 483
 DB 1381 GAAATGACTCAAAAGCTACCAAGCTGGGCGCTCAAAATCTTAAAGCGATATCCGCAATC 1440

REFERENCE 1 (bases 1 to 1464)
AUTHORS Izutsu,H. and Matsumoto,A.
TITLE ANTIGEN POLYPEPTIDE INVOLVED WITH CHLAMYDIA PNEUMONAE, DNA CODING FOR THE SAME, RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT CONTAINING THE VECTOR, AND PRODUCTION OF ANTI-CHLAMYDIA PNEUMONAE Patent: JP 1997009974-A 1 14-JAN-1997;
JOURNAL HITACHI CHEM CO LTD
COMMENT
OS Chlamydia pneumoniae
PN JP 1997009974-A/1
PD 14-JAN-1997
PE 20-SEP-1995 JP 1995242095
PR 20-SEP-1994 JP 94P 224711, 28-APR-1995 JP 95P 106010 PI
PC IZUTSU HIROSHI, MATSUMOTO AKIRA
PC C12N15/09,C07H21/04,C07K14/295,C07K16/12,C12N1/21,C12N15/02, C12P21/08,
PC G01N33/53,G01N33/571//A61K39/118,A61K39/118,A61K39/118, PC (C12N1/21,C12R1:19),
PC (C12P21/08,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FH source 1. 1464
FT /strain='YK41',/organism='Chlamydia pneumoniae' FT
FT mat_peptide 1. 1464
FT /product='antigen peptide'.
FEATURES
source Location/Qualifiers
1. 1464
/organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 470 a 331 c 348 g 315 t
ORIGIN
Alignment Scores:
Pred. No.: 4.05e-117 Length: 1464
Score: 2316.00 Matches: 487
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 97.76% Indels: 0
DB: Gaps: 0
US-09-889-314-2 (1-496) x E12520 (1-1464)
QY 4 MetSerIleSerSerSerSerGlyProAspGlnIleLysAsnIleMetSerGlnValIleu 23
DB 1 ATGTCTATTTCATCTCTTCAGACCTGACAAATCAAAAATATCATGTCTCAAGTTCTG 60
QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGluThrLysGln 43
DB 61 ACATCGACACCCAGCGCGTGGCCCAACAAAGATTAAGCTGTCTGGCAACGAACGAACCA 120
QY 44 IleGlnGlnThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThrIleAlaGly 63
DB 121 ATTCACCAACACAGTCAGGCTAAACACATGAGATGGAAGCGCATGCTGCTATTGCTGCT 180
QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83
DB 181 GCTTCGGAACAAACAAATCTCTCGACATACAAAAACAAAAACAGCTCCACAAACAGGGA 240
QY 84 ValAlaAlaGlySerGluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyValSer 103
DB 241 GTTGCTGCTGGGAAGATCTCTCAGAAATCAAAAGCAGAGTGTCTACTGAGATATCA 300
QY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
DB 301 GGAGCGGCTGCTACTACAGCATCAAAATACGCAACAAATTTGCTATGACAGCTCTATT 360
QY 124 GluGluAlaSerLysSerMetGluSerThrLeuGlnSerLeuGlnSerIleSerAlaAla 143
DB 361 GAAGAGCGGAGCAAAAGTATGAGTACTACTTATGAGTCACTTCAAAACCTCAGCGCGG 420
QY 144 GlnMetLysGluValGlnAlaValAlaIleuSerGlyLysSerSerGlySer 163

DB 421 CAAATGAAGAAGTCGAAGCGGTTGTTGCTGCTCCCTTCACGAGAAAGTTCGGTTCC 480
QY 164 AlaLysLeuGluThrProGlnLeuProLysProGlyValThrProArgSerGluValIle 183
DB 481 GCAAAATGGGAACACTGAGCTCCCAAGCCGGGTGACACCAATATGAGAGTTATC 540
QY 184 GluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThrLysSerAla 203
DB 541 GAATTCGAGCTCGGCTTGGCTAAAGCAATTCAGCATTTGGAGAGAGCCACAAATTCGCC 600
QY 204 LeuSerAsnTyraLaserThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlu 223
DB 601 TTATCTACTACTATGCAAGTACACAGCAGCAAGCAGCAACCAAAATTAAGTCTAGAA 660
QY 224 LysGlnAlaIleLysIleAspLysGluArgGluGluThrGlnGluMetLysAlaAlaGlu 243
DB 661 AAGCAAGCGATAAATAATCGATTAAGAAACGAGAAACAAATACAGATGAAGCGTCGAA 720
QY 244 GlnLysSerLysAspLeuGluGlyThrMetAspThrValAsnThrValMetIleAlaVal 263
DB 721 CAGAACTCTTAAGATCTCGAAGGACAAATGATGCTCATACTGATGATCGCGGTT 780
QY 264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu 283
DB 781 TCTGTGCCATTTACAGTATTATTTCTATTTGCTGCTATTTTTCATCGCGAGCTGCACTC 840
QY 284 AlaGlyLeuAlaAlaGlyAlaAlaValaGlyAlaAlaAlaAlaAlaGlyAlaAlaGlyAla 303
DB 841 GCTGGAGCTCCCTGGGAGCGTGTGAGTGACAGCGGCACTGGAGGTGACGAGGAGCT 900
QY 304 AlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValaGlnAlaValLys 323
DB 901 GCTGCCCAACACGAGTACGACAAACAAATTAACATTAAAGCTGTGTCACAGCGGTGAAA 960
QY 324 GlnAlaValIleThrAlaValaArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLys 343
DB 961 CAACTCTTTTACACAGCTGTCAGACAAAGCATCCCGGCTATTAAGAGCGGTGTCAA 1020
QY 344 SerGlyLysLeuAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363
DB 1021 TCTGGAATTAAGCATTTATCAAAACCTTAGTCAAAACGATGGCAAGCCATTTCCTAAA 1080
QY 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383
DB 1081 GGAATCTTAAGTTTTCGCTTAAGGAGACTCAATGATTTGCGAGAACTTCCCAAGCTTC 1140
QY 384 SerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValaGlyValaVal 403
DB 1141 TCGAAAGTCAATCTGCTCTTACACAGTAATAGGTACAGGTGGGTTGGGTTGTAGTT 1200
QY 404 AlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGluMetGlnGlnAsn 423
DB 1201 GCGCGCGCTCTCTCGGTTAAAGGATTAATGCAATGACGCTCGGAGTGAACAAAC 1260
QY 424 ValAlaGlnPheGlnLysGluValaGlyLysLeuGlnAlaAlaAlaAspMetIleSerMet 443
DB 1261 GTGCTCAATTTTCGAAGAAGTCGAAAACTCGAGGCTCGGCTGATATGATTTTCATG 1320
QY 444 PheThrGlnPheTrpGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyLysSerAsn 463
DB 1321 TTCACCTCAATTTTGGCAACAGCAAGTAATAATTGCTCAAAACAAACAGCGAGTCAAT 1380
QY 464 GluMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyraAlaAlaIle 483
DB 1381 GAATGACTCAAAAGCTTACCAAGCTGGCGCTCAAAATCTTAAGCGTATGCGGCATATC 1440
QY 484 SerGlyValIleAlaGlyAla 490
DB 1441 AGCGAGCCATCGCTGGCGCA 1461
RESULT 13
E12535

Db 1501 AATGAATGACTCAAAAAGCTACCAAGCTGGCGCTCAATCTTAAAGCGTATCCGCA 1560
Oy 483 IleserGlyAlaIleaglYala 490
Db 1561 ATCAGCGGAGCCATCGCTGGCGCA 1584
RESULT 11
ARI22105
LOCUS ARI22105 1464 bp DNA linear PAT 16-MAY-2001
DEFINITION Sequence 3 from patent US 6165478.
ACCESSION ARI22105
VERSION ARI22105.1 GI:14106422
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 1464)
AUTHORS Izutsu,H., Obara,K. and Matsumoto,A.
TITLE DNA encoding Chlamydia pneumoniae antigenic polypeptide
JOURNAL Patent: US 6165478-A 3 26-DEC-2000;
FEATURES
source 1. 1464
/organism="unknown"
BASE COUNT 470 a 331 c 348 g 315 t
ORIGIN
Alignment Scores:
Pred. No.: 4.05e-117 Length: 1464
Score: 2316.00 Matches: 487
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 97.76% Indels: 0
DB: 6 Gaps: 0
US-09-889-314-2 (1-496) x ARI22105 (1-1464)
Oy 4 MetSerIleSerSerSerSerGlyProAspAsnGlnIlyAsnIleMetSerGlnValIleu 23
Db 1 ATGTCTATTTCATCTTCTTCAGAGCCTGCACATCAAAAATATCATCTCTCAAGTTCTG 60
Oy 24 ThrSerThrProGlnGlyAlaProGlnGlnAspIlyLeuSerGlyAsnGluThrIlySer 43
Db 61 ACATGACACCCCGAGGCGTCCCAACAGATAGCTGTGCGACAGCAAGCAAGCAA 120
Oy 44 IleGlnGlnThrArgGlnGlnIlyAsnThrGlnMetGlnSerAspAlaThrIleaglY 63
Db 121 ATACGACCAACGCTGAGGTAAACACTGAGATGGAAGCGATGCCACTATGTCTGCT 180
Oy 64 AlaSerGlyIlyAspIlySerThrSerThrThrIlyThrGluThrAlaProGlnGlnGly 83
Db 181 GCTTCTGGAAAAGACAAAACCTTCGACTACAAAACAGAAACAGCTCCACACAGGGA 240
Oy 84 ValAlaIleaglYIlyGlnSerSerGlnIlyAlaIleaglYAlaAspThrGlyValSer 103
Db 241 GTTGCTGTGGGAAAGAACTCTCAGAAAGTCAAAAGGCGTGTCTGATCTGAGATCA 300
Oy 104 GlYAlaIleAlaIleThrThAlaSerAsnThrAlaThrIlyIleAlaMetGlnThrSer 123
Db 301 GAGGCGGCTGCTACTACACATCAATACTGCAACAAAATTGCTATGAGACCTCTATT 360
Oy 124 GluGlnAlaIleSerIlySerMetGlnSerThrIleuGlnIlySerIleuSerIleuSerAla 143
Db 361 GAAGAGCGGAGCAAAAAGTATGAGTCTACCTTAGAGTCAAGCTCAAGCTTCAGTCCGG 420
Oy 144 GlnMetIlyGlnIlyAlaIleAlaValAlaIleAlaIleuSerGlyIlySerSerGlySer 163
Db 421 CAAATGAAGAAGTCGAACGCTGTGTGTGTCTGCCCTCCAGGGGAAAGTTCCGGCTTC 480
Oy 164 AlaIlyIleuGlnIlyThrProGlnIlyProIlyProIlyValIlyThrProArgSerGlnIly 183
Db 481 GCAAAATTGGAAACACCTGAGCTCCCAAGCCCGGGGTACACCAAGATCAGAGGTTATC 540
Oy 184 GluIleGlyIleuAlaIleuAlaIlyAlaIleGlnIlyThrIleuGlnIlyAlaIlyThrSerAla 203

Db 541 GAAATCGCAGCTCGCCCTGCTAAAGCAATTCAGACATTCGGAGAGCACAAAATCTGCC 600
Oy 204 IeuSerAsnThrAlaSerThrGlnIleaglYAlaAspGlnThrAspIlyIleuGlnIly 223
Db 601 TTATCTAATCTATGACAGTACACAGCACACAGCAAGCAAAATTAATCTAGCTTAGAA 660
Oy 224 IySGlnAlaIleIlyIleAspIlyGlnArgGlnIlyIlyGlnIlyMetIlyAlaIleu 243
Db 661 AAGCAAGCATATAAATATCATTAAGACAGACAGAAAGATACCAAGATGAAGCTCCGGA 720
Oy 244 GlnIlySerIlyAspIleuGlnIlyThrMetAspThrValAsnThrValMetIleAlaVal 263
Db 721 CAGAAGCTAAAGATCTCGAAGGAACAATGATCTGCAATGATGATGATCCGGCTT 780
Oy 264 SerValAlaIleIleThrValIleSerIleValAlaIleIleThrCysGlnIlyIleu 283
Db 781 TCTGTTGCCATTACAGTATTTCTATGTGTGCTGCTATTTTACATGCGGAGCTGACTC 840
Oy 284 AlaGlyIleuAlaIleaglYAlaIleAlaValGlyAlaIleaglYAlaIleaglYAla 303
Db 841 GCTGGACTGCTGCGGAGCTGCTGAGGTGCAAGCGGACGTGGAGTCCACAGAGCT 900
Oy 304 AlaAlaIleAlaThrThrValAlaIleThrGlnIleThrValGlnAlaValAlaIly 323
Db 901 GCTGCCGCAACACGAGTACACACAAATTCAGTTCAAGCTGTGTGCCAAGCGTGAA 960
Oy 324 GlnAlaIleAlaIleThrAlaValArgGlnAlaIleThrAlaIleIleIlyAlaIleVal 343
Db 961 CAAGCTGTTATACAGCTGTCAGACAGGATACCGCGGCTATTAACCGGCTGTCAAA 1020
Oy 344 SerGlyIleIlyAlaIleThrIleIlySerIleuValIlyAlaIleAlaIlyAlaIleSer 363
Db 1021 TCTGGAATTAAGCATTTTCTCAAACTTATGTCAAAAGCATTCGCAAGCCATTTCTAA 1080
Oy 364 GlyIleSerIlyValIleAlaIlySerGlyThrGlnMetIleAlaIlyAsnIleProIlyIleu 383
Db 1081 GGAATCTTAAGGTTTTCGCTAAGGAACTCAATGATTCGAAAGAACTTCCCAAGCTC 1140
Oy 384 SerIlyValIleSerSerIleuThrSerIlyStrpValIleThrValGlyValGlyVal 403
Db 1141 TCGAAAGTATCTGCTCTTACAGTAAATGGGTACAGGTGGGTGGGTGTGTGATT 1200
Oy 404 AlaAlaProAlaIleuGlnIlySerGlyIleMetGlnMetGlnIleuSerGlnMetGln 423
Db 1201 GCGGCGCTGCTCTCGGTAAAGGATTTATGCAAAATGCACTCTCGAGATGCAACAAA 1260
Oy 424 ValAlaGlnPheGlnIlyGlnValGlyIlyIleuGlnAlaIleAlaAspMetIleSerMet 443
Db 1261 GTGCTCAATTCAGAAAGAACTCGGAAAACATGCAAGGCTGCGGTGATGTCTTATG 1320
Oy 444 PheThrGlnPheThrPheGlnIleAlaSerIlyIleAlaSerIlyGlnIlyIlySerAsn 463
Db 1321 TTCACCTCAATTTTGGCAACAGGCAAGTAAATTTGCTCAAAAACAAACAGCGCATTA 1380
Oy 464 GlnMetThrGlnIlyAlaIleThrIlySerIleuGlnIlyAlaGlnIleIleuIlyAlaIle 483
Db 1381 GAAATGACTCAAAAAGCTACCAAGCTGGGCGCTCAATCTTAAAGCGTATGCCGAATC 1440
Oy 484 SerGlyAlaIleIleaglYAla 490
Db 1441 AGCGGAGCCATCGCTGGCGCA 1461
RESULT 12
E12520
LOCUS E12520 1464 bp DNA linear PAT 27-APR-1998
DEFINITION DNA encoding an antigen peptide of Chlamydia pneumoniae.
ACCESSION E12520
VERSION E12520.1 GI:3251353
KEYWORDS JP 1997009974-A/1.
SOURCE unidentified.
ORGANISM unidentified.
unclassified.

polypeptide fused protein.
ACCESSION E16671
VERSION E16671.1 GI:5711354
KEYWORDS JP 1998212298-A/1.
SOURCE unidentified.
ORGANISM unidentified
REFERENCE 1 (bases 1 to 1587)
AUTHORS Iizutsu,H. and Hagimura,T.
TITLE HISTIDINE TAG CHLAMYDIA PNEUMONIAE ANTIGEN POLYPEPTIDE FUSION PROTEIN, ITS PRODUCTION, DNA CODING FOR THE SAME FUSION PROTEIN, RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT CONTAINING THE RECOMBINANT VECTOR, RECOMBINANT BACULOVIRUS, ITS PRODUCTION AND PRODUCTION OF ANTI- CHLAMYDIA PNEUMONIAE ANTIBODY
Patent: JP 1998212298-A 1 11-AUG-1998;
JOURNAL HITACHI CHEM CO LTD
COMMENT OS None
OC Artificial sequences.
PN JP 1998212298-A/1
PD 11-AUG-1998
PF 31-JAN-1997 JP 1997018522
PI IZUTSU HIROSHI, HAGIMURA TOSHIKATSU
PC C07K16/12,C07K19/00,C12N5/10,C12N7/00,C12N15/09,C12P21/02, PC C12P21/08,
PC G01N33/531,G01N33/571//A61K39/395,C07K14/295,(C12N5/10, PC C12R1:91),
PC (C12P21/02,C12R1:91);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key
FH location/Qualifiers
FT source 1..1587
FT location/Qualifiers
1..1587
/organism='Artificial sequences'.
source
BASE COUNT 499 a 363 c 384 g 341 t
ORIGIN
Alignment Scores:
Pred. No.: 3,94e-117 length: 1587
Score: 2317.00 matches: 487
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.80% Mismatches: 0
Query Match: 97.80% Indels: 0
Gaps: 0
US-09-889-314-2 (1-496) x E16671 (1-1587)
QY 3 AsnMetSerIleSerSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnVal 22
Db ::
Db 121 TCCATGCTATTTCATCTCTTCAGACGCTGACAAATCAAAAAAATATCATGTCTCAAGTT 180
QY 23 LeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGluThrLys 42
Db ::
Db 181 CTGACATCGACACCCAGGCGCTGCCCAACAAGATTAAGTGTCTGGCAACGAAACGAAG 240
QY 43 GlnIleGlnGlnThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThrIleAla 62
Db ::
Db 241 CAAATTCAGCAACACGCTCAGGGTAAACACCTGAGATGGAAGCCGATGCCACTATTGCT 300
QY 63 GlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGlnGln 82
Db ::
Db 301 GGGCGCTTGTGAAAAGCAAACTCTCTCGACTACAAAAAAGAGAAAGAGCTCCACACAG 360
QY 83 GlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyVal 102
Db ::
Db 361 GGGATTGCTGCTGGGAAAGATCTCTGAAAGATCAAAAGGACAGGTGCTGATTACGAGTA 420
QY 103 SerGlyAlaAlaAlaThrThrAlaThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSer 122

Db 421 TCAGAGCGCGCTGCTACTACAGCATCAATATCTCAACAAAAAATTGCTATGACAGACTCT 480
QY 123 IleGluGluAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerIleSerAla 142
Db ::
Db 481 ATTGAAAGGCGAGCAAAAGATGAGAGTCTTACCTTAAGTCACTTCAAACTCAGTGC 540
QY 143 AlaGlnMetLysGluValGlnAlaValAlaValAlaAlaLeuSerGlyLysSerSerGly 162
Db ::
Db 541 GCGCAATTCGAAAGAGTCGAAGCGGTTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 600
QY 163 SerAlaLysLeuGluThrProGlnLeuProLysProGlyValThrProArgSerGluVal 182
Db TCCGCAAAATTCGAAACACCTGAGCTCCCAAGCCGGGGAGACACCAAGATCAGAGCTT 660
QY 183 IleGluIleGlyLeuAlaLeuAlaValAlaIleGlnThrLeuGlyGluAlaThrLysSer 202
Db 661 ATCGAAATTCGAGCTCGGCTGTCTTAAGCAATTCAGCATTTGGAGAGCAACAAAATCT 720
QY 203 AlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeu 222
Db GCTTATCTACTATGCAAGTACCAAGCACAGCAGACCAACAAATTAACCTAGTCTA 780
QY 223 GluLysGlnAlaIleLysIleAspLysGluArgGluGluLutyrGlnGluMetLysAlaAla 242
Db 781 GAAAGACGACGATAAATATCGAATTAAGAACGAAACGAAATACCAAGATGAAGCTGCC 840
QY 243 GluGlnLysSerLysAspLeuGluGlyThrMetAspThrValAsnThrValMetIleAla 262
Db 841 GAACGAGAGTCTAAAGATCTCGAAGGAAACAATGATCTGTCAATCTGTGATGATCTCGC 900
QY 263 ValSerValAlaIleThrThrAlaIleSerIleValAlaAlaIlePheThrGlySerGlyAlaGly 282
Db 901 GTTCTCTTGCCATTACAGTATTATTCTATTGTCTCTCTATTTTTACATCGGGAGCTGA 960
QY 283 LeuAlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaAlaGlyAlaAlaGly 302
Db 961 CTGCTGAGACTCGCTGGGGAGCTGCTGTAGGTGCACGCGGAGCTGAGGTGCACGAGGA 1020
QY 303 AlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaVal 322
Db 1021 GCTGCTCGCCGCAACGACGAGTACCAACAATTCACGTTCAAGCTGTGTTCACAGCGCTG 1080
QY 323 LysGlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaAlaVal 342
Db 1081 AAACAGACTGTATCACAGCTGTACAGACGATCACCGCGGTATTAAGGCGCTGC 1140
QY 343 LysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSer 362
Db 1141 AAATCTGCAATTAAGCATTTATCAAAACCTTAGTCAAAAGCGATTCGCAAAAGCATTTCT 1200
QY 363 LysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLys 382
Db 1201 AAAGCAATCTCTAAAGTTTCGTAAGGAACTCAATTAATGATGGAGAACTTCCCAAG 1260
QY 383 LeuSerLysValIleSerSerLeuThrSerLysThrValThrValGlyValGlyValAl 402
Db 1261 CTCTCGAAAGTCACTGCTGCTTACCAAGTAAGAGGGTCAAGGTGGGTGGGGTGTGA 1320
QY 403 ValAlaAlaProAlaLeuGlnGlyLysGlyIleMetGlnMetGlnLeuSerGluMetGlnGln 422
Db 1321 GTTCCGGCGCGCTGCTCGGTAAAGGATTAATGCAATTCAGCTCTCGGAGATGCAACAA 1380
QY 423 AsnValAlaGlnPheGlnLysGluValGlyLysLeuGlnAlaAlaAlaAspMetIleSer 442
Db 1381 AACGTCCTCAATTTTCAGAAAGAGTCGGAATACTGCAAGCTGGCGCTGATATGATTTCT 1440
QY 443 MetPheThrGlnPheThrProGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyGluSer 462
Db 1441 ATGTTCACTCAATTTTGGCAACAGCAAGTAATTAATTCCTCAAAAACAAACAGCGACTCT 1500
QY 463 AsnGluMetThrGlnLysAlaThrLysLeuGlyValAlaGlnIleLeuLysAlaThrAlaAla 482

```
PN JP 1998212298-A/7
PD 11-AUG-1998
PF 31-JAN-1997 JP 1997018522
PI IZUNSU HIROSHI, HAKIMARA TOSHIKATSU
PC C07K16/12,C07K19/00,C12N5/10,C12N7/00,C12N15/09,C12P21/02, PC
C12P21/08,
PC G01N33/531,G01N33/571//A61K39/395,C07K14/295,(C12N5/10, PC
C12R1.91),
PC (C12P21/02,C12R1.91);
CC strandedness: Double;
CC topology: Circular;
FH Key Location/Qualifiers
FT source 1..11099
FT FT /organism='Artificial sequences' FT CDS
FT 4135..5601
FT /product='Chlamydia pneumoniae antigen FT
FT polyptide'
FT misc_feature 7..4029
FT /note='DNA sequences derived from FT
FT baculovirus' 5635..8467
FT misc_feature /note='DNA sequences derived from FT
FT baculovirus' 4030..4092.
FT promoter
FT Location/Qualifiers
FT 1..11099
FT /organism='unidentified'
FT /db_xref='taxon:32644'
BASE COUNT 3075 a 2453 c 2525 g 3046 t
ORIGIN

Alignment Scores:
Pred. No.: 2,816-116 Length: 11099
Score: 2320.00 Matches: 490
Percent Similarity: 99.60% Conservative: 2
Best Local Similarity: 99.19% Mismatches: 1
Query Match: 97.93% Indels: 1
De: 6 Gaps: 0

US-09-889-314-2 (1-496) x E16677 (1-11099)
OY 3 AsnMetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnVal 22
DB 4132 TCCATGTCATTTCATCTCTTCAGACCTCGACATCAAAAATATTCATGCTCAGATT 4191
OY 23 LeuThrSerThrProGlnGlyValProGlnAspLysLeuSerGlyAsnGlnThrLys 42
DB 4192 CTGACATCGACACCCCGAGGCGTGGCCCAACAGATTAAGCTGTCTGGCAACGAAAG 4251
OY 43 GlnIleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspIleThrIleAla 62
DB 4252 CAATACAGCAACACGTGAGGTAAACACGTGAGTGAAGGCGATGCTACTATTGCT 4311
OY 63 GlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGln 82
DB 4312 GGTGCTTCGGAAGAACAAACTTCTCGACACAAAAACAAACAGCTCCACAAACAG 4371
OY 83 GlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyVal 102
DB 4372 GGAGTGTGCTGGTGGAAAGATCTCGAAGAGCAAAAGGCGAGTGTACTGTGAGTA 4431
OY 103 SerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSer 122
DB 4432 TCGAGAGCGGCTCTACTACATCAAAATACGCAACAAAAATGCTATGACAGACCTCT 4491
OY 123 IleGlnGlnAlaSerLysSerMetGluSerThrLeuGlnLysLeuGlnSerLeuSerAla 142
DB 4492 ATTGAAGAGCGGAGCAAAATATGAGTCTACCTTAGAGTCACTTCAAAAGCTCAGTGCC 4551
OY 143 AlaGlnMetLysGluValGluAlaValAlaValAlaLeuSerGlyLysSerSerGly 162
DB 4552 GCCCAATATGAAGAGTGAAGCGGTGTGTGCTGCCCTTCAGGGAAGATTCGGGC 4611
```

```
OY 163 SerAlaLysLeuGluThrProGlnLeuProLysProGlyValThrProArgSerGluVal 182
DB 4612 TCCGCAAAATTTGGAACACCTAGCTCCCAAGCCGGGGGTACACCAAGATCAAGGTT 4671
OY 183 IleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThrLysSer 202
DB 4672 ATCGAAATCGGACCTCGCGCTTCTTAAGCAATTCAGACATTGGGAGAACCCAAACTT 4731
OY 203 AlaLeuSerAspThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeu 222
DB 4732 GCCATTATCTAATGCAAGTACACAGCAAGCAAGCAAGCAACAAATTAACCTTAGTCTA 4791
OY 223 GluLysGlnAlaIleLysIleAspLysGluArgGlnGluGluThrGlnLysLeuAlaAla 242
DB 4792 GAAAGCAAGCATATAAATATCATAAAGACAGAGAAGATACCAAGATGAAGGCTGCC 4851
OY 243 GluGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAla 262
DB 4852 GAACAGAAAGTCTAAAGATCTCGAAGCAACAAATGATCTGTCAAATCTGTGATCGCG 4911
OY 263 ValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGly 282
DB 4912 GTTCTGTTGCCATTACATTAATTCTATTGTGCTGCTATTATTACATGCGGAGCTGGA 4971
OY 283 LeuAlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaAlaAlaGlyAlaAlaGly 302
DB 4972 CTCGCGAGACGCTGCTCGGAGCTGCTGTAAGTGCAGCGGAGCTGAGTGCAGACAGA 5031
OY 303 AlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaVal 322
DB 5032 GCTGCTGCCGCAACACCGGTAGCAACAAATTAAGTTCAGTTCAGTTCCTTCCAAAGGGTG 5091
OY 323 LysGlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaAlaVal 342
DB 5092 AAACAAGCGTTATATCACCTGTCAGACCAAGCATACCGCGCTATTAAGGGGTGTC 5151
OY 343 LysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSer 362
DB 5152 AATCTGGAATAAAGCATTTATCAAAACTTATGTCAAAGCGATTGCCAAAGCATTTCT 5211
OY 363 LysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLys 382
DB 5212 AAAGGATCTCTTAAGGTTTTCGCTAAGGCAACCAATATGCGAAGAACTTCCCAAG 5271
OY 383 LeuSerLysValIleSerSerLeuThrSerLysThrValThrValGlyAlaValAla 402
DB 5272 CTCTCGAAAGTCATCTGCTCTTACCAAGTAATGGGTCAAGTTGGGTTGGTGTGA 5331
OY 403 ValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnGln 422
DB 5332 GTTGGCGCGCTCTCTCGTAAAGGATTATGCAATGACAGCTTCGAGATGCAACAA 5391
OY 423 AsnValAlaGlnPheGlnLysGluValAlaGlyLysLeuGlnAlaAlaAlaAspMetIleSer 442
DB 5392 AACGTGCTCAATTTCAAGAAAGTCCGAAACCTCAGGCTGCGCTGTATGATGATTTCT 5451
OY 443 MetPheThrGlnPheThrGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyLysSer 462
DB 5452 ATGTTCACTCAATTTTGGCAACAGCAAGTAATATGCTCAAAACAAACAGCGAGTCT 5511
OY 463 AsnGlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaThrAlaAla 482
DB 5512 AATGAATGACTCAAAAGAGTACCAAGCTGGGCGCTCAAACTCTTAAGGATGTCGCGA 5571
OY 483 IleSerGlyAlaIleAlaGlyAla-HisLysThrAsnAsn 495
DB 5572 ATCAGCGGAGCCATCGCTGCGCAGCATTAAGCTGAAT 5611
RESULT 10
E16671 1587 bp DNA linear part 28-JUL-1999
LOCUS E16671
DEFINITION DNA encoding histidine tag - Chlamydia pneumoniae antigen
```

```

promoter          4030..4092
FT               /note='Baculovirus polyhedrin gene promoter'
FT CDS           4135..5601
FT              /product='Chlamydia pneumoniae 53kDa-antigen
FT              peptide'
FT              5635..8467
FT              /note='Baculovirus-derived sequence'.
FEATURES
   source         location/Qualifiers
                   1..11099
                   /organism='unidentified'
                   /db_xref='taxon:32644'
BASE COUNT      3075 a 2453 c 2525 g 3046 t
ORIGIN
Alignment Scores:
Pred. No.:       2,81e-116      Length:      11099
Score:           2320.00        Matches:      490
Percent Similarity: 99.60%      Conservative: 2
Best Local Similarity: 99.19%   Mismatches: 1
Query Match:     97.93%        Indels:      1
DB:              6            Gaps:        0
US-09-889-314-2 (1-496) x E16636 (1-11099)

QY 3 AsnMetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnVal 22
Db 4132 TCCATGCTCAATTTCATCTCTTCAGACCTGCATCAAAAAAATATCATGCTCAAGTT 4191

QY 23 LeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLys 42
Db 4192 CTGACATGACACCCAGGCGGCTGCCACAAAGATAGCTGTGGGACAGAAAGCAAG 4251

QY 43 GlnIleGlnInThrArgGlnGlyLysAsnThrGlnMetGlnLysSerAspAlaThrIleLea 62
Db 4252 CAANACAGCAACACGTCAGGCTAAACACCTGATGCAAGCGATGCGCATATTGCT 4311

QY 63 GlyAlaSerGlyLysAspLysThrSerSerThrLysThrGlnThrAlaProGlnGln 82
Db 4312 GGTGCTCTGGAAAAAGACAAACTCCCTCGACTACAAAAACAGAAACAGCTCCACACAG 4371

QY 83 GlyValAlaAlaGlnGlyLysGlnSerSerGlnLysAlaGlyAlaAspThrGlyVal 102
Db 4372 GGAGTGTGCTGGGAAAGAAATCCCTCAGAAAGTCAAAAGGAGCGCTGCAATCTGAGCTA 4431

QY 103 SerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSer 122
Db 4432 TCAGAGAGCGGCTGCTACTACAGCATCAAAATCTGCACAAAAAAATTTGCTATGACAGCTCT 4491

QY 123 IleGlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAla 142
Db 4492 ATTGAAGAGCGGCAAAAGATAGAGTCTACCTTAGAGTCACTTCAAAAGCCTCAGTGCCT 4551

QY 143 AlaGlnMetLysGlnValGlnAlaValValAlaAlaAlaLeuSerGlyLysSerSerGly 162
Db 4552 GCGCAAAATGAAAGAGTGAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 4611

QY 163 SerAlaLysLeuGlnThrProGlnLeuProLysProGlnValThrProArgSerGlnVal 182
Db 4612 TCCGCAAAATGGAACACCTCGAGCTCCCAAGCCGGGGTGACACCAAGATCAGAGGTT 4671

QY 183 IleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlyAlaThrLysSer 202
Db 4672 ATCGAAATTCGAGACTCGCGCTGTGTAAGCAATTCAGACATTTGGGAGAAAGCAAAATCT 4731

QY 203 AlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeu 222
Db 4732 GCCCTATCTACTATGCAAGTACCAAGCAACACCAAGCAAAATTAATTAATTAATTAATTAAT 4791

QY 223 GlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGln 242
Db 4792 GAAAAAGCAAGGATAAATATGATTAAGAACGACAAATATTCAGAAAGATGAAGGCTGCC 4851

QY 243 GlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAla 262

```

```

Db 4852 GACACGAGACTCTAAAGATCTCGAAGGAAACATGATGATCTCATACTGATGATGCGG 4911
QY 263 ValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGly 282
Db 4912 GTTCTGTGTGGCATACAGATTATTTCTATGTTGCTGCTATTTTACATGGCGAGCTGGA 4971
QY 283 LeuAlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAlaAlaAlaGlyAlaAlaGly 302
Db 4972 CTCGCTGACTCGCTGCGGAGAGCTGCTAGCTGCAGCGGAGCGAGCTGCGAGCGAGCA 5031
QY 303 AlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaVal 322
Db 5032 GCTGCTGCCGACACACCGGCTGACCAACAAATTCAGTTCAAGCTGTTGCTCAAGCGGTG 5091
QY 323 LysGlnAlaValAlaIleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaAlaVal 342
Db 5092 AAACAAAGCTGTATACAGCTGTGCAGACAAAGCATCACCGCGGCTATTAAGCGGCTGTC 5151
QY 343 LysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSer 362
Db 5152 AAATCTGGAATTAAGCAATTTATCAAAACTTTAGTCAAAAGCATTTGCCAAAGCATTTCT 5211
QY 363 LysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLys 382
Db 5212 AAAGAAATCTTAAGGTTTTCCTTAAGGAACTCAAAATGATGGCAAGAACTTCCCAAG 5271
QY 383 LeuSerLysValIleSerSerLeuThrSerLysThrValThrValGlyValGlyValAla 402
Db 5272 CTCCTGAAAGTCATCTGCTCTTACAGCAATTAATGGGTACAGGTTGGGGGTGGGTTGTA 5331
QY 403 ValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLysSerGlnMetGlnGln 422
Db 5332 GTTGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 5391
QY 423 AsnValAlaGlnPheGlnLysGlnValGlyLysLeuGlnAlaAlaAlaAspMetIleSer 442
Db 5392 AACGTCGCTCAATTTTCAGAAAGAGTGGAAAGATGCGAGCTGCGGCTGATATCATTTCT 5451
QY 443 MetPheThrGlnPheThrPheGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyLysSer 462
Db 5452 ATGTTCACTCAATTTTGGCAACAGCAAGTAAATTTGCTTCAAAACAAACAGCGAGTCT 5511
QY 463 AsnGlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaThrAlaAla 482
Db 5512 AATGAATGACTCAAAAGAGTCAACAGCTGGCGCTCAAAATCTTAAAGCGTATGCGCA 5571
QY 483 IleSerGlyAlaIleAlaGlyAla-HisLysThrAsnAsn 495
Db 5572 ATCAGCGGAGCCATCGCTGGCGGACAGCATTAAGCTAGAAAT 5611

RESULT 9
LOCUS      E16677
DEFINITION antigen polypeptide-coding DNA.
ACCESSION E16677.1
VERSION   1
KEYWORDS  JP 1998212298-A/7.
SOURCE    unclassified.
ORGANISM  unclassified.
REFERENCE 1 (bases 1 to 11099)
AUTHORS   Itzutsu,H. and Hagiwara,T.
TITLE     HISTIDINE TAG CHLAMYDIA PNEUMONIAE ANTIGEN POLYPEPTIDE FUSION
          PROTEIN, ITS PRODUCTION, DNA CODING FOR THE SAME FUSION PROTEIN,
          RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT CONTAINING THE
          RECOMBINANT VECTOR, RECOMBINANT BACULOVIRUS, ITS PRODUCTION AND
          PRODUCTION OF ANT-CHLAMYDIA PNEUMONIAE ANTIBODY
          Patent: JP 1998212298-A 7 11-AUG-1998;
          HITACHI CHEM CO LTD
COMMENT   OS None
          OC Artificial sequences.

```



```

|||||
Db 16236 ATGTGCTGCTCTGAAAAAGACAAACTTCCGACCTACCAAAAACAGAAACGCTCCA 16177
QY 81 GlnGlnGlnValAlaAlaGlyLysSerSerGlnSerGlnLysAlaGlyAlaAspThr 100
Db 16176 CAACAGGAGCTCTGCTGGGAAAGAACTCTCAGAAAGCTCAAAAGGCGAGTGGTGAATCT 16117
QY 101 GlnValSerGlnValAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120
Db 16116 GAGATATCAGAGAGCGGTGCTACTACGCAATCAAAATCTGCACAAAATTTGCTATGCTCAG 16057
QY 121 ThrSerIleGlnGlnLysSerSerMetGlnSerThrLeuGlnSerLeuGlnSerLeu 140
Db 16056 ACCTATATGAAGAGGAGCAAAAGATGAGTCTACCTTAAAGTCACTTCAAAAGCTC 15997
QY 141 SerAlaAlaGlnMetLysGlnValGlnAlaValAlaAlaLeuSerGlyLysSer 160
Db 15996 AGTGCAGCGCAAAATGAAAGAGTCGAAAGCGGTGTTGCTGCTCCCTCAGGGAAGT 15937
QY 161 SerGlnSerAlaLysLeuGlnThrProGlnLeuProLysProGlnValThrProArgSer 180
Db 15936 TCGGGTTCCCAAAATTTGGAAACACCTGAGCTCCCAAGCCGGGGTGAACACCAAGATCA 15877
QY 181 GlnValIleGlnIleGlnLysLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlnAlaThr 200
Db 15876 GAGGTATTCGAAATCGGAGCTCGGCTTGCTAAAGCAATTCAGACATTTGGAGAACCCACA 15817
QY 201 LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
Db 15816 AAATTCGCTTATCTAATCTATGACATGACAAAGCACACAGCCCAAAACAAATTAACCTA 15757
QY 221 GlnLeuGlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnIleGlnGlnMetLys 240
Db 15756 GGTCTAGAAAAAGACAGCATTAATAATCGATTAAGAACGACGAAATACCAAGAGATGAG 15697
QY 241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMet 260
Db 15696 GCTGCCGAACAGAAAGTCTTAAGATCTCGAAGGAACATGATCTCACTACTGCTGATG 15637
QY 261 IleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGly 280
Db 15636 ATGCGGCTTCTGCTGCAATCTACAGTTATGCTTCTGCTCTATTTTAACTGGGGA 15577
QY 281 AlaGlnLysLeuAlaGlnLysLeuAlaGlnAlaValAlaGlnAlaAlaGlnGlyAla 300
Db 15576 GCTGGAGCTCGCTGACTCGCTGGGAGCTGCTGTAGCTCAGCGGAGCTGGAGGTGCA 15517
QY 301 AlaGlnValAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGln 320
Db 15516 GCAGGAGCTGCTCGCCGACACGAGTAGCAACAAATTTACAGTCAAGCTGTGTCCAA 15457
QY 321 AlaValLysGlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAla 340
Db 15456 GCGGTAACCAACACTTATACACAGCTGTCAGACAAAGCGATCACCGGCTTAAAAAACG 15397
QY 341 AlaValLysSerGlnLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAla 360
Db 15396 GCTGCTAAATCTGCAATTAAGCATTTATCAAACTTTACTCAAGCGATTTGCCAAAGCC 15337
QY 361 IleSerLysGlnLysSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPhe 380
Db 15336 ATTTCTAAAGGATCTCTAAGTTTCGTAAGGAACTCAAAATGATTCGAGAAACTTC 15277
QY 381 ProLysLeuSerLysValIleSerSerLeuThrSerLysTryValThrValGlyValGly 400
Db 15276 CCCAACTCTCGAAAGTCATCTGCTCTTACACAGTAAATGGGTTCACGAGTTGGGTTGG 15217
QY 401 ValValValAlaAlaProAlaLeuGlnLysGlyIleMetGlnMetGlnLeuSerGlnMet 420
Db 15216 GTTGTAGTTGGCGGCTGCTCGGTAAGGGAATTAATGCAATTCGAGCTCGGAGATG 15157
QY 421 GlnGlnAsnValAlaGlnPheGlnLysGlnValAlaGlyLysLeuGlnAlaAlaAspMet 440
|||||

```

```

Db 15156 CAACAAACGTCGCTCAATTTCAGAAAGAGTCGAAACACTGCAGGCTGGCGCTGATATG 15097
QY 441 IleSerMetPheThrGlnPheThrGlnGlnAlaSerLysIleAlaSerLysGlnThrGly 460
Db 15096 ATTTCTATGTTCACTCAATTTTGGCAACAGCAAGTAAATTTGCTCAAAACAAACAGCG 15037
QY 461 GlnSerAsnGlnMetThrGlnLysAlaThrLysLeuGlnLysAlaGlnIleLeuLysAlaThr 480
Db 15036 GAGCTAATGAAATGATCTCAAAAAGCTAACAGCTGGGCGCTCAAAATCTTAAAGCTAT 14977
QY 481 AlaAlaIleSerGlnAlaIleAlaGlyAla-HisLysThrAsnAsnPhe 496
Db 14976 GCGGCATTCAGCGGAGCGCATCGCTGGCGAGCATAAACCAATTAATTTT 14928
RESULT 7
E16678
LOCUS
DEFINITION
All sequences of plasmid pHRCPM53 containing DNA encoding histidin
tag - Chlamydia pneumoniae antigen polypeptide fused protein.
ACCESSION
E16678.1 GI:5711361
VERSION
JP 1998212298 -A/8.
KEYWORDS
unidentified.
SOURCE
ORGANISM
unclassified.
REFERENCE
1 (bases 1 to 9595)
AUTHORS
Izutsu,H. and Hagiwara,T.
TITLE
HISTIDINE TAG CHLAMYDIA PNEUMONIAE ANTIGEN POLYPEPTIDE FUSION
PROTEIN, ITS PRODUCTION, DNA CODING FOR THE SAME FUSION PROTEIN,
RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT CONTAINING THE
RECOMBINANT VECTOR, RECOMBINANT BACULOVIRUS, ITS PRODUCTION AND
PRODUCTION OF ANTI- CHLAMYDIA PNEUMONIAE ANTIBODY
Patent: JP 1998212298-A 8 11-AUG-1998;
HITACHI CHEM CO LTD
COMMENT
OS None
OC Artificial sequences.
PN JP 1998212298-A/8
PD 11-AUG-1998
PE 31-JAN-1997 JP 1997018522
PI IZUTSU HIROSHI, HAGIWARA YOSHIKATSU
PC C07K16/12,C07K19/00,C12N5/10,C12N7/00,C12N15/09,C12P21/02, PC
C12P21/08,
PC G01N33/531,G01N33/571//A61K39/395,C07K14/295,(C12N5/10, PC
C12R1:91),
PC (C12P21/02,C12R1:91);
CC strandedness: Double;
CC topology: Circular;
CC hypothetical: No;
CC anti-sense: No;
FH Key
FH Location/Qualifiers
FT source 1..9595
FT /organism="Artificial sequences" FT CDS
FT 2206..3795
FT /product="Histidin tag - Chlamydia pneumoniae
FT antigen
FT polypeptide fused protein"
FT misc_feature 1..2128
FT /note="DNA sequences derived from FT
FT baculovirus" 3868..6480
FT misc_feature /note="DNA sequences derived from FT
FT baculovirus" 2129..2191.
FT promoter 1..9595
FT Location/Qualifiers
FT 1..9595
FT /organism="unidentified"
FT /db_xref="taxon:32644"
BASE COUNT 2676 a 2114 c 2191 g 2614 t
ORIGIN
Alignment Scores: 1.12e-116 length: 9595
Pred. No.:

```

```
gene
CDS
LDD"
complement(2558. .3607)
/gene="CPj0799"
complement(2558. .3607)
/gene="CPj0799"
/codon_start=1
/transl_table=11
/product="hypothetical protein"
/protein_id="BAAG9007.1"
/db_xref="GI:8979173"
/translation="MAAIKQILRSMGLSSSLMMVLESLYSLSGYCYITDREPDEHS
SSAVKMDHMKTTLSRLSNKRAKAAVSGTATTVGFKIDTWSRTYAVRWYWGKEL
PTSSWVKSKATGSISSDGSIIAGIYENELSSQFAVTMKNEMYLPTWAVQSKAYGI
SSDSVIYGSADAMSRFEAVKMGHEAOVLPGVAVKANSVANSANGSIYVGSODA
SGIIYAVKMGENTITHTGLTGYSIAKAVNSNGVIYGRSEYTPGFEVHACHKNGVM
SDLTIGGSIYSAKGVSAATGKVIYGMSTTANGKLHAFYVGGRMIDLEIYSWKECAN
AVSIDGETIVGVSE"
complement(3867. .5153)
/gene="eno"
complement(3867. .5153)
/gene="eno"
/codon_start=1
/transl_table=11
/product="enolase"
/protein_id="BAAG9008.1"
/db_xref="GI:8979174"
/translation="MFEAVIADIQARELIDSRGPTLVKVTSTSGVGEARVBSGAS
TGRKEALEFRDTSFRIOGKVLQAVKVKELFPLVGCSEYEDSLIDSLMDSGDS
PKNETLGNAILIGVSLATAHAAATLRPLRYLGCFCASLPCPMNMLINGMADN
GLGQEFMRIRPGASIKAAVNGADVHTLTKLHERGLSTGVDEGGFAPNLASNE
EALDELILAIEAGFTPGKDISLADCAASFEYNKGTGYDRHYEEOIALISNCDR
YPISIDEGLAEEEDYDQWALLTEVLEGEKVOIYGDLLFTVNPDELLEGISNGANSVL
KPNQIGLTETVVAIKLAOMAGYTTIISRGETTDTIADLAFAFNAGQIKTGSLSR
SERAKVYRMLEIEELGSEAFITDSNFTSEDSSE"
complement(5345. .7318)
/gene="uvrB"
complement(5345. .7318)
/gene="uvrB"
/codon_start=1
/transl_table=11
/product="exonuclease ABC subunit B"
/protein_id="BAAG9009.1"
/db_xref="GI:8979175"
/translation="MTFOLHAPFAPCGDOPPAIARLSAGVRNOVSGVLLGTSGKT
FTIANVYANVLPITYLAHNKTLAOLYQEFREFPNNAVEYFISYDYQPEATAR
SDTIEKSLILNDELIDKLRLSATRSILERRDTLYSVSCITGTSPPNTYSMALVE
VGKEYPRNITLAQLVKMHYQASPIQRSAFERGSDVIDIPAYESELARLEFLNDL
TSIEYSDPLTMIIPKESVPSATLYPGSHYIPEAIRQAIRTOELERMAFFDRPI
EKDRIFHTTIDIEIKIEGCKIENYSRHTGAPPGAPPTCLDYPEDFLITIDE
SHOITLPOIRAMVRGDSQSLVEYGFRLPSAFDMRRLPYEEAOKYFRKVIYATPG
DVEYESGHTVVOOITIRRTGIPDMPEIRPAPGVDLLEETRLRLSKRHKIIVIST
TKKLAEBAAGFLSELEIPAYLHSGIETAEKRTQITLDRSGVIDVLIGVNLREGLD
PEVSLVALIDADKGEFLKSTSLIQFCGRARNINGKVIETADQKTRSIETLRGTER
RROITLDYKNKHNIYKPIIKAIFAFNPLQTSKSPKESPKESPESRSLKEDLEQIKKE
ALMRRAKEFFNEAKYRDAMQACEQLLVLF"
complement(7327. .8361)
/gene="trpS"
complement(7327. .8361)
/gene="trpS"
/codon_start=1
/transl_table=11
/product="tryptophanyl tRNA synthetase"
/protein_id="BAAG9010.1"
/db_xref="GI:8979176"
/translation="MNKKRVLTGDRPPGKHLHGWVGSINKRELONSPEYDCFFII
ADLHTLTTKIRKEEVLVDNHIYEVLAWLVSIGDPTSIILYLOSAIPEIYELHLFS
MLISINRWGIPSLKDMARNASIEBGSISYGLIGPIQSADILAKQFYPVGDNE
AHVELTRDIARNFNRLYGQVPEPEVLQGLSTIGVIGOGKMSKANNALYLSQDA
TITEKVRMYDNPRIKATTPGVRGEGNLFYIHDFNHKDEVEEFKARYOGCICKDI
EVKARIAELIHLFLPIKERSESEFLSKLALQNLVEDGTHKMKREYAKATMEVHDKG
FSHKWRSLK"
complement(8370. .8924)
/gene="CPj0803"
complement(8370. .8924)
gene
CDS
complement(8370. .8924)
/gene="CPj0803"
/codon_start=1
/transl_table=11
/product="Ct584 hypothetical protein"
/protein_id="BAAG9011.1"
/db_xref="GI:8979177"
/translation="MAAKTKULEEDNVFLLLEGNLRKIRFAPPIGYTTFREONVFN
CANGQELANFEPELITNGKLTQELAPQOKAASLTAEPMPIRVADDIHERGFIN
FITSMDLTQGERCIPLNRLARVDGEFLMTDVQNTCHLIHLLARLLEAOKNPVGEK
NLQETQEEITSLKHNFDLTALQ"
complement(8928. .9680)
/gene="gp6D"
complement(8928. .9680)
/gene="gp6D"
/codon_start=1
/transl_table=11
/product="CHLrR plasmid paralog"
/protein_id="BAAG9012.1"
/db_xref="GI:8979178"
/translation="MGNKITLIESFRKKNPTPKKALARKMEGDPFLAVLSNPTL
SKKEQRLHLQHTNREQITEEDLTQCLTSLAEVKIHHQSVLLHERITKRDLL
KSYREGAFSSWLLLYGNRQTPYNFLVYELTFLPEPLKTEMKMPQAVYTLASRQ
GPOEKKEEIIIRNYRGEKSELDRIRKEFPLVETDCRKTSPVKQALMLTRGSQILTK
CTSLSDQIILEKLIKLEKYSMLPFDTKV"
complement(9685. .10452)
/gene="mind"
complement(9685. .10452)
/gene="mind"
/codon_start=1
/transl_table=11
/product="chromosome partitioning ATPase"
/protein_id="BAAG9013.1"
/db_xref="GI:8979179"
/translation="MKTIAVNSFKGTAKTSTTHLGAALQYHQARVLLIDPDAQAN
LTSGLDPCDYDSLAVVLQGEKEIOEYIRPIODQDLIPADTLEIRIEVSGNLAAD
RYSHERLKYVLGSVDKDYVIIDTPPLCMTESALIAAYALICATPEYVSGYLE
RLAGITQGISARHPLTILIGVALSEFVNCNGKNSAFAEIHHFTPEKLLNTIRIDITV
SEAHGKRPVATSPSARASSEDYFNLRELLLRDI"
complement(10485. .12392)
/gene="lhrs"
complement(10485. .12392)
/gene="lhrs"
/codon_start=1
/transl_table=11
/product="threonyl tRNA synthetase"
/protein_id="BAAG9014.1"
/db_xref="GI:8979180"
/translation="MIQVTCQKNVEVLGGTTAAELAKQLKNSHQFIGVLIINERPOL
Alignment Scores:
Pred. No.: 1,36e-115 Length: 325865
Score: 2340.00 Matches: 496
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 0
Query Match: 98.78% Indels: 1
DB: 1 Gaps: 0
US-09-889-314-2 (1-496) x AP002548 (1-325865)
QY 1 AAPTThrAsnMetSerIleSerSerSerGlyProAspAsnGlnInlysAsnIleMetSer 20
|||||
DB 16416 GATACAAACATGTCATTTTCATCTTCTTACGACCGACGAAATCAAAAAATATCATGCTC 16357
QY 21 GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlu 40
|||||
DB 16356 CAAGTTCGACATCGACACCCAGGCGCGCCCAACAAGATAAGCTGCTCGCAACGAA 16297
QY 41 ThrLysGlnIleGlnGlnIleThrArgGlnGlyLysAspThrGluMetGluSerAspAlaThr 60
|||||
DB 16296 ACGAAGCAATACAGCAACACGTCAGGGTAAACACAGTGAAGGCAATGCGACT 16237
QY 61 IleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaPro 80
```

Db 4744 TCGGGTCCGCAAAATTTGAAACACACTGAGCTCCCAAGCCGGGGTGCACCAAGATCA 4803
 QY 181 GluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThr 200
 Db 4804 GAGGTATTCAGAAATCGGACTCGCGCTTGCTAAACCAATTGCAGACTGGGAGAGCCACA 4863
 QY 201 LysSerAlaLeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
 Db 4864 AATATTCCTTATCTACTATCGAAGTACACAGACACAGACAGCCAAACAAATAACTA 4923
 QY 221 GlyLeuGluLysGlnAlaIleLysIleAspLysGluArgGluTyrGlnGluMetLys 240
 Db 4924 GGCTTAAGAAAAGCAAGCGATAAATATGATTAACAGAGCAAGATATCCAAAGATGAAG 4983
 QY 241 AlaAlaGluGlnLysSerLysAspLeuGlyThrMetAspThrValAsnThrValMet 260
 Db 4984 GCTGCCCAAGAGAGTCTAAAGATCTGAAGGAACAATGGTACTGTCAAATACGTATG 5043
 QY 261 IleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGly 280
 Db 5044 ATCGCGGTTTCTGTGGCATTACAGTTATTTCTATTGCTGCTATTTTACATGCGGA 5103
 QY 281 AlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAlaAlaAlaGlyGlyAla 300
 Db 5104 GCTGACTCGCTGACTCGCTGCGGAGCTGCTAGTGCAGCGGAGCTGAGGTGCA 5163
 QY 301 AlaGlyValAlaAlaAlaIleThrValAlaIleThrGlnIleThrValGlnAlaValGln 320
 Db 5164 GCAGGACTCTCGCCGACCCAGCCAGTACCAACAAATTACAGTTCAAGCTGTGTCGCA 5223
 QY 321 AlaValLysGlnAlaValIleThrValAlaValArgGlnAlaIleThrAlaAlaIleLysAla 340
 Db 5224 CGCGTGAACACAGCTGTATACACAGCTGTACAGACACAGATCCCGCGCTATATAAGCC 5283
 QY 341 AlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAla 360
 Db 5284 GCCTCAATCTGGAATTAAGCATTTATCAAACTTAGCAAAAGGATGCCAAAGCC 5343
 QY 361 IleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPhe 380
 Db 5344 ATTTCTAAAGAAATCTTAAGTTTTCGTAAGGAACTCAATGATTGCGAAGACTTC 5403
 QY 381 ProLysLeuSerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValGly 400
 Db 5404 CCCAAGCTCTCGAAGCATCTCTCTCTTAACCAAGTAAATGGGTACAGGTTGGGTTGG 5463
 QY 401 ValValValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGluMet 420
 Db 5464 GTTGTACTTGGCGGCGCTCGCTCGGTAAGGAGATTATGCAAAATGCAAGCTCTCGAGATG 5523
 QY 421 GlnGlnAsnValAlaGlnPheGlnLysGlyValGlyLysLeuGlnAlaAlaAspMet 440
 Db 5524 CAACAAAACGCTCAATTTCAAGAAAGAGTCCGAAAACCTGCAGCGCTCGATATG 5583
 QY 441 IleSerMetPheTrpGlnPheTrpGlnAlaSerLysIleAlaSerLysGlnThrGly 460
 Db 5584 ATTTCTATGTTCACTCAATTTTGGCAACAGCAAGTAAATTTGCTCAAAAACAACAGCC 5643
 QY 461 GluSerAsnGluMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyr 480
 Db 5644 GAGCTTAATGAAATGACTCAAAAAGTACCAAGCTGGGCGCTCAAAATCCTTAAGCGTAT 5703
 QY 481 AlaAlaIleSerGlyAlaIleAlaGlyAla-HisLysThrAsnAsnPhe 496
 Db 5704 GCCGCAATCAGCGGAGCATCGTGGCGCAGCATAAACCAATATTTT 5752
 RESULT 6
 AP002548/c 325865 bp DNA linear BCT 25-MAY-2002
 LOCUS Chlamydomophila pneumoniae J138 genomic DNA, complete sequence,
 DEFINITION section 4/4.
 ACCESSION AP002548 AB033788 AB033789 AB033790 AB033791 AB033818 AB036083
 AB036084 AB036085 AB036086 AB036087 AB036088 AB036089 AB038350

AB038351 AB038352 BA000008
 AP002548.2 GI:10176694
 Chlamydomophila pneumoniae J138 (strain:J138) DNA.
 Chlamydomophila pneumoniae J138
 Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 1
 Shirai,M., Hirakawa,H., Ouchi,K., Tabuchi,M., Kishi,F., Kimoto,M.,
 Takeuchi,A., Nishida,J., Shibata,K., Fujiwara,R., Yoneda,H.,
 Matsushima,H., Tanaka,C., Furukawa,S., Miura,K., Nakazawa,A.,
 Ishii,K., Shiba,T., Hattori,M., Kuhara,S. and Nakazawa,T.
 Comparison of outer membrane protein genes omp and pmp in the whole
 genome sequences of Chlamydia pneumoniae isolates from Japan and
 the United States
 J. Infect. Dis. 181 Suppl 3, S524-S527 (2000)
 20298986
 2
 Shirai,M., Hirakawa,H., Kimoto,M., Tabuchi,M., Kishi,F., Ouchi,K.,
 Shiba,T., Ishii,K., Hattori,M., Kuhara,S. and Nakazawa,T.
 Comparison of whole genome sequences of Chlamydia pneumoniae J138
 from Japan and CWL029 from USA
 Nucleic Acids Res. 28 (12), 2311-2314 (2000)
 20330349
 3 (bases 1 to 325865)
 Shirai,M.
 Direct Submission
 Submitted (04-JUL-2000) Mutsunori Shirai, Yamaguchi University
 School of Medicine, Department of Microbiology; 1-1-1
 Minamikogushi, Ube, Yamaguchi 755-8505, Japan
 E-mail:mshirai@po.cc.yamaguchi-u.ac.jp, Tel:81-836-22-2227,
 Fax:81-836-22-2415)
 On or before Sep 15, 2000 this sequence version replaced
 gi:6172302, gi:6172304, gi:6172306, gi:6172308, gi:6172400,
 gi:6635182, gi:6635184, gi:6635186, gi:6635188, gi:6635190,
 gi:6635192, gi:6635194, gi:8547442, gi:8547447, gi:8547451,
 gi:8979170.
 Location/Qualifiers
 1..325865
 /organism="Chlamydomophila pneumoniae J138"
 /strain="J138"
 /db_xref="taxon:138677"
 /note="synonym:Chlamydia pneumoniae (strain J138)"
 221..1318
 /gene="CpJ0797"
 221..1318
 /gene="CpJ0797"
 /codon_start=1
 /transl_table=1
 /product="hypothetical protein"
 /protein_id="BAA99005.1"
 /db_xref="GI:8979171"
 /translation="MSKRIKIVDLHLITCLTFRGVLCALASINIGVSTSOEFPYOKSI
 EDMKGYTFDLELLSKGMSFAHISNGSRIYVAGSAGGSVAYIWSHILRKHGT
 LGESASABEGISKRDGEVYVWSDTRREGYTKAFVFDGDMDLTGLATYASVAGSGD
 GSIILVGSATFARGEDYGMQVGVKWKGIKQLKLLPQGLSEANALISDGTIVVGKE
 ISRNHLYAVKWNNAVYSLIGTIGSVASAAISANKVITGVSTNNNGEHNHMHDE
 TMHDLGTLGGSPSVATGVSADGAIYFVSFKVGEIHAFYTAGEMHDLTTLGGEAR
 VFDISBGNIDIGSIKTDGAERAYLFHHK"
 1467..2480
 /gene="CpJ0798"
 1467..2480
 /gene="CpJ0798"
 /codon_start=1
 /transl_table=1
 /product="hypothetical protein"
 /protein_id="BAA99006.1"
 /db_xref="GI:8979172"
 /translation="MKTKCCONRSGIVPSVYLFTVTTQTLRAGHFIDTIGSLYSM
 ARGVSGDGRVYVTEGSAKRIYDGEFTLEGLVPSSEALVFRASIDGVTIGISDD
 PSORAVKWNVALVDLGIIFSEGMQSFAGVSSDGKTIIVGCLYSDDTETNFAVWDEG
 MVVLPIPEDRHSCAMDASDGSVIVODAMGSEIARAVYWKDGEHLLSNIPGAKRS
 SAHVSXKDSFIYGEIFSENEVHAIFYHNGVLIKIDIGLGGDSVATGVSADGKVIYV
 HSRITDGEYAFKIVDGRMIDLDTIGSASAFARVAGSDGKTIYKRETELGECHAFIY

```
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF3834.1"
/db_xref="GI:7189975"
/translation="MNKPKTKKAVQKAPVKVPEBSDEAIIQIQLAVSLDYKE
LPAAETFASLDKNOINSIIAALSGTLESLHLEELTQGLFPAQDANFAKELSSVH
GLKNJTVNKKQVKGAE"
4273..5739
/gene="CP1062"
4273..5739
/gene="CP1062"
4273..5739
/feature="conserved hypothetical protein; identified by
glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF3835.1"
/db_xref="GI:7189976"
/translation="MSISSSGPDNOKNIMSOVLSTPQGVPOQDKLSGNETKOIOT
ROGKTEEMESDATTIAGSGDKSTSTKTETAPQGVAAKRESSEKSGAGADTVSGA
AATASNTATRIAMOTSTIEPAKSMESTLESTQISAAOMKEVAVVVAALSGSGS
AKTEPELPKRGVPRSEVIEIGLAKAIQTLGERTKSALSNASTOAOADQTKRLG
LEKQAIKIDKEEYOEKMAEQKSDLEGTWDTVTVMIAVSAITVISTVAAIFTC
GAGLAGAAGAAGAAAGAAATTAATVQVAVQAVQAVITAVQATITAA
IKAAVSGIKAFIKTLVKAIAKISGISKSTPAKOTIAKNEPKLSKISLSLTKWV
TVGVGVVAADALGKIMOMOLEMOQVAVOFOKEVKGIAAADMISMFTQWQASK
IASKOTGSENMTOKATKLGAIILKAIYAIISATAGAA"
5772..7106
/gene="CP1063"
5772..7106
/gene="CP1063"
/feature="conserved hypothetical protein; identified by
glimmer2; putative"
/codon_start=1
/transl_table=11
/product="conserved hypothetical protein"
/protein_id="AAF3836.1"
/db_xref="GI:7189977"
/translation="MTSGVSGSSSDPTLAAQIQAOSQKAGNAGSHDKNVTKQAO
AEVAGGPEPDIODASQSTCKEATSTKSKSEKSKSSSTSVASSETAT
KAOVQGPGLKRONNTDPSLTPPEKQTINGILKGMGTLLGLVMTLMANAGESW
KAEQSONQAIKRSQVESAPAIGEAIKROANHQAQATEQAKSLISGLVNIAGFVS
GAGIFSAKGAATSAKSAFPAKETSAGAGASALTSASSSVQOTMASTAKAATTA
ASSAGSAATKAANLTDMAAASAKMASAGASAGAFGEVLNKPNNSEKSGRCMNV
VKTGARVASPAGNALSSMOMLHGHTAIVELISAGQGTIEVAHHORLAGQAEQ
AEVLKOMSSVYGGQAGQAGQOQBOAMQSFNTALQTLNINADSOQTTSATFN"
complement(7177..8094)
/gene="CP1064"
complement(7177..8094)
/feature="membrane protein, putative; identified by match to
PFAM protein family HMM PF00892"
/codon_start=1
/transl_table=11
/product="membrane protein, putative"
/protein_id="AAF73728.1"
/db_xref="GI:8163547"
/translation="MHMSFLVFLTAFTWSSSFALSKLVMAASATIPATGARMTAGA
ITALAMFRGCGVGISSKIFLYIVLLALGFLVLTNIFETIGQSLSSSTCFIYGLSP
LMSALFSYIOLKEKVTYKLVGLSLGLVSYICLYLPFGGSGDQSPWQIGPELLIL
GAASLASFGWTLRLQIERKOSTLSTVAINAYAMLIAGLMSISMSAVEPRLPVODIS
OFLVATLAVVISNLICYNLYAKRLKRYSTFLSPCNVIMPLYSFGYGIILGKGVGS
LGIVLAVFMVACRLIYHEFRQGIYS"
8288..10195
/gene="CP1065"
8288..10195
/feature="similar to GB:M36593 SP:P18256 PID:143764
PID:1565235 GB:AL009126; identified by sequence
similarity; putative"
/codon_start=1
/transl_table=11
/product="threonyl-tRNA synthetase"
```

```
/protein_id="AAF3837.1"
/db_xref="GI:7189978"
/translation="MIOYTCDOKNYEVLEGTAAELAKOLKNSHOFIGVLINERPLD
STHNEGDTLVLTSEDPBGRGRTFLHTSAHLIAQVILKLPADITPTIPVDHGRYTD
FANLISSEDEPLIEDTVKQIVDEKLAISRFYVGGQQLAQPONPKRTTELRLPE
NEELISYSGEFPEDLCRPHLPSTAHVAFAKVLRTSAAYWRQDPRESLVRYGFSF
TSKELRHLLEQIEEAKKRDHRYLGAKLDFSGQESSPCMPFHPBGMIVMDALIRYWK
OLHTAGYKEILTPOLMNRDLMVEGSHMDNYKANNYTLQIDDEDYAIKPMNPGCMY
YKTRLHSYKEPRLRYAEVGHVROEASGALSGLVRAVAFHODAHVPLTPBQVEETL
NITQVSTLYTGFTGLEIYHLELSTPREKOTIGDSDIMELATDANALVOSGPTIYRP
GEGATYPRKIDIHVADAIDQRTWQCGTIDLMFLPRELEIETTAQGTSPVMLRAL
FGSIERFGLILENFGRFPLMLSPQVARIIVADRIIPRAELEAMKRLGLVYLD
DSSESSEKRIKRNQNMQVNYMITLGDHEINENVLAVRTDRNRYINDVSVERPLNILE
EKNSLSTALL"
10228..10995
/gene="CP1066"
10228..10995
/feature="similar to GB:D26185 SP:P37522 PID:467381
PID:580906 GB:AL009126; identified by sequence similarity;
putative"
/codon_start=1
/transl_table=11
/product="Para family protein"
/protein_id="AAF3838.1"
/db_xref="GI:7189979"
/translation="KKTAVANSFKGCTAKTSTTTLHGAALAOYHOARVLLIDPDQAN
LTSGLGDPDCDLSIAVVLQGEKEIOEVIRIOPDQLDIPADTWLERIEVSGNLAD

Alignment Scores:
Pred. No.: 4,74e-118 Length: 15455
Score: 2356.00 Matches: 496
Percent Similarity: 99.808 Conservative: 0
Best Local Similarity: 99.808 Mismatches: 0
Query Match: 99.458 Indels: 1
DB: Gaps: 0

US-09-889-314-2 (1-496) x AE002263 (1-15455)

QY 1 AspThrAsmMetSerIleSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20
|||||
DB 4264 GATVCAAAACATGCTCTTTCATCTCTTCAGACCTGCACATCAAAAATATCATGCTCT 4323
QY 21 GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlu 40
|||||
DB 4324 CAAGTCTGACATCGACACCCAGGCGTGCCTCCCAACAAGATAGCTGTCTGCAACGAA 4383
QY 41 ThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThr 60
|||||
DB 4384 ACGAACAATAATACGCAACACGCTGAGGGTAACACATGATGGAAGACATGCGACAT 4443
QY 61 IleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaPro 80
|||||
DB 4444 ATTGCTGGTGCCTTCGAAAAAGACAAAACCTCTCTGACATCAAAAACAGAACGCTCCA 4503
QY 81 GlnGlnGlyValAlaIleGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThr 100
|||||
DB 4504 CAACAGAGAGTTCCTCGGAAAGAAATCTCAGAAAGCTAAAAGCAGGCTGTAATCT 4563
QY 101 GlnValSerGlyAlaIleAlaIleThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120
|||||
DB 4564 GAGATGATCAGAGACGCGCTGCTACTACAGATCAAAATACGCAACAAAATATGCTATGACG 4623
QY 121 ThrSerIleGlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeu 140
|||||
DB 4624 ACCCTCATTTGAAAGAGCGACCAAAAGTATGAGATCACTTAAGTACATCTTCAAAACCTTC 4683
QY 141 SerAlaIleGlnMetLysGlnValGlnAlaValAlaValAlaIleAlaLeuSerGlyLysSer 160
|||||
DB 4684 AGTGCCTGCGCAATGAAAGACGAGCGGTGTTGTTGCTGCTCCCTCAGGGAAGAAAGT 4743
QY 161 SerGlySerAlaLysLeuGlnThrProGlnLeuProLysProGlnIleValThrProArgSer 180
|||||
```

Oy	281	AAGCTGTTAAACGTGGTGAATCAACTGCATTTCATAAAGCAGCGCATCCGCCGCTTAAGAACC	300
Dd	1067	GCTGAGCTTGCGTAGCTGCTGGAGCTGCTGAGTGGACAAGCGCACGTGAGAGTCCA	1126
Oy	301	AAGCTGTTAAACGTGGTGAATCAACTGCATTTCATAAAGCAGCGCATCCGCCGCTTAAGAACC	320
Dd	1127	GCAGAGCTGCTGCGCGCACCGGTAGAACCAAATTACAGTTCAAGCTGTTCCAA	1186
Oy	321	AlaValLysGlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaLeuSala	340
Dd	1187	GCAGTGAACAAGCTGTTATCACAGCTGTCAGACAAAGCATCCGCCGCTTAAGAACC	1246
Oy	341	AlaValLysSerGlyIleLeuSalPheIleIleuTrpLeuValLysAlaIleAlaLysSa	360
Dd	1247	GCTGTCNAATCTGGAGATTAAGACATTTATCAAACTTATAGTCAAGCAATGGCCAAAAGC	1306
Oy	361	IleSerLySGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPe	380
Dd	1307	ATTCTTAAGAAGAACTCTAAGGTTTTTGGCTTAAGGAAGCACTCAAAATGATTTGCCAAGAATTC	1366
Oy	381	ProLysIleuSerLysValIleSerSerLeuThrSerLysTyTPvalThrValGlyValGly	400
Dd	1367	CCTCAAGCTCTCGAAAGTATCATCTGCTCTTACCAAGTAATGGCTCACCGGTGGGTTGGG	1426
Oy	401	ValValIleAlaAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGluNet	420
Dd	1427	GTTTGATGTTGGCGGCGCTGCTCTCGGTAAAGGATTAATGCAATGCAGCTCTCGGAGATG	1486
Oy	421	GInGlnAsnValAlaGlnPheGlnIleLysGluValGlyLysLeuGlnAlaAlaAspmet	440
Dd	1487	CAACAAAAACGTGCTCATATTTCAGAAAGAAAGTCGAAAACCTGCAGCGCTGATATG	1546
Oy	441	IleSerMetPheThrGlnPheTrpGlnGlnAlaSerLysIleAlaSerLysGlnThrGly	460
Dd	1547	ATTCTTAAGTTCACCAATTTTGGCAACAGCAATAAATTGGCTCTCAAAACAAACAGCG	1606
Oy	461	GluSerAsnGluMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlatyr	480
Dd	1607	GAGCTAATGAATGATCACTCAAAAACCTCCAAAGCTGGCGCTCAAAATCTTTAAACGATAT	1666
Oy	481	AlaAlaIleSerGlyAlaIleAlaGlyAla-HisLysThrAsnaSnpe	496
Dd	1667	GCCGCAATCAGAGGAGCCATCGCTGGCGCAGCATTAACCAATATATTTT	1715
RESULT 5			
AE002263			
LOCUS	AE002263 1545 bp DNA linear BCT 30-MAY-2000		
DEFINITION	Chlamydophila pneumoniae AR39, section 89 of 94 of the complete genome.		
ACCESSION	AE002263 AE002161		
VERSION	AE002263.2 GI:8163546		
KEYWORDS			
SOURCE			
ORGANISM	Chlamydophila pneumoniae AR39. Bacteria; Chlamydiales; Chlamydiaceae; Chlamydophila.		
REFERENCE	1 (bases 1 to 1545)		
AUTHORS	Read,T.D., Brunham,R.C., Shen,C., Gill,S.R., Heidelberg,J.F., White,O., Hickey,E.K., Peterson.J., Umayam,L.A., Utterback,T., Berry,K., Bass,S., Linher,K., Weidman,U., Khouri,H., Craven,B., Bowman,C., Dodson,R., Gwin,M., Nelson,W., Deboy,R., Kolonay,J., McClarity,G., Salizberg,S.L., Eisen,J. and Fraser,C.M. Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39 Nucleic Acids Res. 28 (6), 1397-1406 (2000)		
TITLE			
JOURNAL MEDIANE PUBLISHED REFERENCE AUTHORS	JOURNAL MEDIANE PUBLISHED REFERENCE AUTHORS		
2 (bases 1 to 1545) Read,T.D., Brunham,R.C., Shen,C., Gill,S.R., Heidelberg,J.F., White,O., Hickey,E.K., Peterson.J., Umayam,L.A., Utterback,T., Berry,K., Bass,S., Linher,K., Weidman,U., Khouri,H., Craven,B., Bowman,C., Dodson,R., Gwin,M., Nelson,W., Deboy,R., Kolonay,J., McClarity,G., Salizberg,S.L., Eisen,J. and Fraser,C.M.			

TITLE Direct Submission
 JOURNAL Submitted (01-MAR-2000) The Institute for Genomic Research, 9712
 COMMENT Medical Center Dr, Rockville, MD 20850, USA
 FEATURES On Jun 1, 2000 this sequence version replaced gi:7189971.
 source Location/Qualifiers
 1..15455
 /organism="Chlamydomonas reinhardtii"
 /strain="AR39"
 /db_xref="taxon:115711"
 /note="synonym: Chlamydomonas reinhardtii"
 /complement(121..1188)
 /gene="Cp1058"
 /complement(121..1188)
 /gene="Cp1058"
 /note="similar to GB:3372642; identified by sequence
 similarity; putative"
 /codon_start=1
 /transl_table=1
 /product="proline dipeptidase"
 /protein_id="AAP38833.1"
 /db_xref="GI:7189972"
 /translation="MSHRLRLAQRALSEHNDAIVKESDIALYFLHDAIAGILLI
 GGQVFEFFVYKMDKDLSDIQRVLTPTQPVNDAISYQKRYKIGDSASTVYH
 KFAQVLYPCIMEDLECFETKIRISKSEEFIRRMQEAALGASGYDVLTLRGIEIEH
 KEVQRRLARAFEAAGAGSPSPPIIAGGSAFPHSIIPTDRLKKGIDVILDIGVLN
 GXCSDMTKMTALGCTPHPKILSEPVYVNAQKRAMALCKEYIMQDIDAEAVRVREHHH
 LDITFTHGIGHGVRHNIHEPCSPRGSQVKLESMTITVEPGVIFPGIGIRIDTLIC
 IDKKNFSLTARPISELVCL"
 /complement(1193..2935)
 /gene="Cp1059"
 /complement(1193..2935)
 /gene="Cp1059"
 /note="similar to SP:P23367 GB:Z11831 PID:42067 PID:537011
 GB:000096; identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=1
 /product="DNA mismatch repair protein mutL"
 /protein_id="AAP38832.1"
 /db_xref="GI:7189973"
 /translation="MSTRRPQLDPLPTINOIAGAEVYIENSVSVKELIENSIDAGAD
 EIELEITGGGGAIIIRDNGCGFAEDPIALQHRATSKIREFDISLNSFGRGRA
 PLASIASIKRGIOSIEGDEGVRVIVGGDVSCEPCAROLGTIVINSVNVRRR
 GFSQMSQDRKIGIRKILNRLITRANIGSMWISGSHHEITQAKGQGOERAYVMGSH
 FMQALITDKRANSVRIVGVGSPSPFRPTROGKITINDRPIESLTSKKVQDAYLL
 LPLHRYRVEVLYLPLPSSWCDVFNHQRKEARILKEELVGDCKEAVIETLACPAII
 LCRHLIEESDSVLPFLRMLETSDVEESVFDONLEAVSSEDSLEQETYSRGSI
 PKSMDMIVSDVETFLSLGKRVLAEDLEGVHIIFTAAKRLHFLSLIMOENSHYQSO
 QALLPIRLQVTPPEAFEPFHHGRTLDLGIEISQVGCYSIESPTVIGEEELKEMK
 LLIAGSGTDINSEALTLAKKELTLQATPESKHAYDVSMLKILMSVGRDEKFDGDA
 RIRLLIIDSPEMG"
 /3158..3853
 /gene="Cp1060"
 /3158..3853
 /gene="Cp1060"
 /note="similar to GB:M26405 SP:P21207 PID:155451;
 identified by sequence similarity; putative"
 /codon_start=1
 /transl_table=1
 /product="type III secretion chaperone SycD"
 /protein_id="AAP38833.1"
 /db_xref="GI:7189974"
 /translation="MSKRPSPRNANOPQPSASFNKTKTRSLAEILAQKAKADLEQV
 HPVTEIEIKKALGNIEGLSNGIDLQDILGLSDYLLLEETIVATYVTSQGYNEAVG
 LFQILAAQAPQNTKYMGLSSCYHQHLHYNNAAGCFIAPACPDNPIIPYVYADSLLL
 KLOAPDESNPLDVTMDICGNPFIKLKERQQLMKQSIKQMGATGKKAFTKPKAPAK
 SKTNTNKGSKKR"
 /3874..4236
 /gene="Cp1061"
 /3874..4236
 /gene="Cp1061"
 /note="conserved hypothetical protein; identified by
 GLIMMER2; putative"
 /codon_start=1

Db 1247 GCTGTCAATCTGGAAATTAAGACATTTATATCAAACTTTAGTCAAGCGATTGCCAAAGCC 1306
Qy 361 ILESerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAspHe 380
Db 1307 ATTTCTAAGGAATCTTAAGGTTTTCCTAAGGAACTCAAAATGATTGCGAAGACCTTC 1366
Qy 381 ProLysLeuSerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValGly 400
Db 1367 CCCAAGCTCTCGAAGATCATCTCGCTCTTACACATAATGGGTACAGGTTGGGTTGG 1426
Qy 401 ValValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMet 420
Db 1427 GTTGTAGTTGGCGCCCTCTCGTAAAGGATTATGCAAAATGCACCTCTCGAGATG 1486
Qy 421 GlnGlnAsnValAlaGlnPheGlnLysGlnLysGlyValGlyLysLeuGlnAlaAlaAspMet 440
Db 1487 CAACAAAACGTGCTCAATTTACAGAAAGATCGGAAAACATGCAAGCTGCGGCTATATG 1546
Qy 441 ILESerMetPheThrGlnPheTrpGlnGlnAlaSerLysIleAlaSerLysGlnThrGly 460
Db 1547 ATTTCTATGTTCACTCATTTTGGCAACAGCACTAAATATGCTCAAAACAAACAGGC 1606
Qy 461 GlnSerAsnGlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyr 480
Db 1607 GAGTCTAATGAATGACTCAAAAGCTACCAAGCTGCGGCTCAATCTTAAGCGTAT 1666
Qy 481 AlaAlaIleSerGlyAlaIleAlaGlyAla-HisLysThrAsnAspHe 496
Db 1667 GCCGCAATCAGCGGAGCCATCGCTGCGCAGCATATAAACCAATATATTT 1715

RESULT 4
EL6674
LOCUS EL6674 1939 bp DNA linear PAT 28-JUL-1999
DEFINITION ~~GenBank~~ Clamydia pneumoniae antigen polypeptide.
ACCESSION EL6674
VERSION EL6674.1 GI:5711357
KEYWORDS JP 1998212298-A/4.
SOURCE unclassified.
ORGANISM unclassified.
REFERENCE unclassified.
1 (bases 1 to 1939)
AUTHORS Izutsu,H. and Hagiwara,T.
TITLE HISTIDINE TAG CHLAMYDIA PNEUMONIAE ANTIGEN POLYPEPTIDE FUSION PROTEIN, ITS PRODUCTION, DNA CODING FOR THE SAME FUSION PROTEIN, RECOMBINANT VECTOR CONTAINING THE DNA, TRANSFORMANT CONTAINING THE RECOMBINANT VECTOR, RECOMBINANT BACULOVIRUS, ITS PRODUCTION AND PRODUCTION OF ANTI- CHLAMYDIA PNEUMONIAE ANTIBODY
PATENT: JP 1998212298-A 4 11-AUG-1998;
HITACHI CHEM CO LTD
JOURNAL OS Chlamydia pneumoniae
PN JP 1998212298-A/4
PD 11-AUG-1998
PI 31-JAN-1997 JP 1997018522
PI 1ZUTSU HIROSHI, HAGIWARA TOSHIKATSU
PC C07K16/12,C07K19/00,C12N5/10,C12N7/00,C12N15/09,C12P21/02, PC C12P21/08,
PC G01N33/531,G01N33/571//A61K39/395,C07K14/295,(C12N5/10, PC C12N1:91),
PC (C12P21/02,C12N1:91):
CC strandness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH key Location/Qualifiers
FT source 1. 1939
FT /organism='Chlamydia pneumoniae' FT
FT /strain='YK-41' /clone='53-35'
FT 5'UTR 1..235
FT CDS 236..1702
FT /product='Chlamydia pneumoniae antigen FT
FT polypeptide'

FEATURES FT 3'UTR 1703..1939.
source 1..1939
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 631 a 433 c 450 g 425 t
ORIGIN
Alignment Scores:
Pred. No.: 3,91e-119 Length: 1939
Score: 2356.00 Matches: 496
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 0
Query Match: 99.45% Indels: 1
DB: Gaps: 0
US-09-889-314-2 (1-496) x EL6674 (1-1939)
Qy 1 AspThrAsnMetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20
Db 227 GATCAAAACATGTCATATTTCATCTTCTCAGGACCTGCACAAATCAAAAATATCATGTCT 286
Qy 21 GlnValLeuThrSerThrProGlnGlyValProGlnIleAspLysLeuSerGlyAsnGln 40
Db 287 CAAGTTCGACATCGACACCCGAGGCGTCCCAACAAAGATTAACCTGCTGGCAACGAA 346
Qy 41 ThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThr 60
Db 347 ACGAAGCAAAATACACCAACACGTCAGGTTAAAAACACTGAGATGGAAGCGATCCACT 406
Qy 61 ILeAlaGlyAlaSerGlyLysAspLysThrSerSerThrLysThrGluThrAlaPro 80
Db 407 ATTGCTGGTGGCTTGGGAAAACAAACTCTCTGACATCAAAAAACAGAAACAGCTCCA 466
Qy 81 GlnGlnGlyValAlaAlaGlyLysGlnSerSerGlnLysAlaGlyAlaAspThr 100
Db 467 CAACAGGAGGAGTGTGCTGGGAAAGATCTTCGAAAGCAAAAGCAGGTCTGATACT 526
Qy 101 GlyValSerGlyAlaAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120
Db 527 GGAGTATCAGGAGCGGCTGCTACTACAGCATCAATCTGCACAAAATTCTTATGCA 586
Qy 121 ThrSerIleGlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeu 140
Db 587 ACCTCATATTGAAGAGCGGCAAAAGTATGAGTCTACTTAGAGTCACTTCAAAAGCCTCA 646
Qy 141 SerAlaAlaGlnMetLysGlnValGlnAlaValValAlaAlaAlaLeuSerGlyLysSer 160
Db 647 AGTGGCGGCAAAATGAAGAAAGTCAAGCGGTTGTGTGCTGCCCTTCACGGGAAAAGT 706
Qy 161 SerGlySerAlaLysLeuGlnLthrProGlnLthrProLysProGlnValLthrProArgSer 180
Db 707 TCGGGTTCGCAAAATGGAAACACTGAGCTCCCAAGCCCGGGGTGACACCAAGATCA 766
Qy 181 GlnValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnIleThrLeuGlyGlnAlaThr 200
Db 767 GAGGTATATGAATATGCAATCGCGCTGTCTAAAGCAATTCACACATTTGGGAAAGCACA 826
Qy 201 LysSerAlaLeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
Db 827 AAATCTGCTTATCTTAACATGCAAGTACACAGCAACAGCAACCAAAATTAACCTTA 886
Qy 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGlnLthrGlnLthrGlnLysGlnLys 240
Db 887 GGTCTAAGAAACGACGATTAATAATCGATTAAGACGAGGAAGATACCAAGATGAAG 946
Qy 241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMet 260
Db 947 GCTCGCAACAGACAGTCAAGATCTCGAAGAACATGATCTCAATCTGATGATG 1006
Qy 261 ILeAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrGly 280
Db 1007 ATCGGGTTTGTGTGCCATACAGTATTTCTATTGTGTGCTATTTTCTACATCCGA 1066

DB 2814 CCCAGCTCTCGAAGTCATCTGCTCTTACCAAGTAATAGGCTGACGGTTGGGTTGGG 2755
QY 401 ValValValAlaAlaProAlaLeuGlyIleMetGlnMetGlnLeuSerGlnMet 420
DB 2754 GTTAGTTCGGCGCGCTGCTCGGTAAGGGATTATGCAATTCGACGCTCTCGAATG 2695
QY 421 GlnGlnAsnValAlaGlnPheGlnIleGlyValGlyIleGlyLeuGlnAlaAlaAspMet 440
DB 2694 CAACAAACGCTGCTCAATTCAGAAAGAGTCGAAACATGCAAGCTGCGCTGATATG 2635
QY 441 IleSerMetPheThrGlnPheThrGlnIleAlaSerIleAlaSerIleGlnIleGly 460
DB 2634 ATTTCATGTTCTCACTCAATTTTGGCAACAGCAAGTAATAATGCTCAAAACAGG 2575
QY 461 GluSerAsnGlnMetThrGlnIleValAlaThrIleGlyValAlaGlnIleLeuValAlaThr 480
DB 2574 GAGTCATTAAGTAATGACTCAAAAGCTACCAACACTGCGGCTCAATTCCTTAAGCTAT 2515
QY 481 AlaAlaIleSerGlyAlaIleAlaIleAlaIleHisIleThrAsnAsnPhe 496
DB 2514 GCGCAATCAGCGAGCAGCATCGCTGCGCACATAAAACCAATATTTT 2467
RESULT 3
E16639
LOCUS E16639 1939 bp DNA linear PAT 28-JUL-1999
DEFINITION Chlamydia pneumoniae gene for 53kDa-antigen peptide.
ACCESSION E16639
VERSION E16639.1 GI:5711322
KEYWORDS JP 1998210978-A/5.
SOURCE unclassified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1939)
AUTHORS Izutsu, H. and Hagihara, T.
TITLE RECOMBINANT VECTOR AND TRANSFORMANT CONTAINING THE SAME, AND
CHLAMYDIA PNEUMONIAE ANTIGEN POLYPEPTIDE
PATENT: JP 1998210978-A 5 11-AUG-1998;
HITACHI CHEM CO LTD
COMMENT OS Chlamydia pneumoniae
PN JP 1998210978-A/5
PD 11-AUG-1998
PF 31-JAN-1997 JP 1997018523
PI 12UTSU HIROSHI, HAGIMARA TOSHIKATSU
PC C12N5/09, C07H21/04, C07K14/295, C12N1/21, C12N5/10, C12N7/00, PC
C12P21/02//
PC A61K39/00, A61K39/118, C12Q1/68, (C12N5/09, C12R1:01), (C12N1/21,
PC C12R1:19),
PC (C12N5/10, C12R1:91), (C12N7/00, C12R1:92), (C12P21/02, C12R1:91);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1. 1939
FT /organism='Chlamydia pneumoniae' FT
FT /strain='YK-41',
FT /tissue_type='elementary bodies' FT CDS
FT 236..1702 /product='53kDa-antigen peptide'.
FEATURES
source Location/Qualifiers
1..1939
/organism="unclassified"
/db_xref="taxon:32644"
BASE COUNT 631 a 433 c 450 g 425 t
ORIGIN
Alignment Scores:
Pred. No.: 3.91e-119 Length: 1939
Score: 2356.00 Matches: 496
Percent Similarity: 99.80% Conservative: 0
Best Local Similarity: 99.80% Mismatches: 0
Query Match: 99.45% Indels: 1
DB: 6 Gaps: 0

US-09-889-314-2 (1-496) x E16639 (1-1939)
QY 1 AspThrAsnMetSerIleSerSerSerSerGlyProAspAsnGlnIleMetSer 20
DB 227 GATACCAACATGCTATTTTCATCTTCTTCAGACCTGACAAATCAAAAAATATCATCTCT 286
QY 21 GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspIleSerGlyAsnGlu 40
DB 287 CAACTTCTGACATCGACACCCAGGCGTCCCAACAAATATAGCTGTGGCAACGAA 346
QY 41 ThrIleGlnIleGlnGlnThrArgGlnGlyIleAsnThrIleMetGlnSerAspAlaThr 60
DB 347 ACAAGCAATATACAGCAACACGTCAGGGTAAAAACACTGATGGAACGATGCGACT 406
QY 61 AlaAlaGlyAlaSerGlyIleAspIleThrSerSerThrThrIleThrAlaPro 80
DB 407 ATGCTGGTCTCTTGAAAAAGCAAAACTCTCCAGCTCAAAAAACAGAACGCTCCA 466
QY 81 GlnGlnGlyValAlaAlaGlyIleGlyIleGlyIleSerSerGlnIleGlyAlaAspThr 100
DB 467 CAACAGGAGTTCTGCTGGGAAAAAGATCTCGAAAAAGTCAAAAAGCGAGTGTGATACT 526
QY 101 GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrIleAlaMetGln 120
DB 527 GAGTATACAGAGCGCGCTGCTACTACAGCATCAATATCTGCAACAAAAATGCTATGACG 586
QY 121 ThrSerIleGlnIleAlaSerIleSerMetGlnSerThrIleGlnIleSerIleu 140
DB 587 ACCCTATTGAAAGAGCGAGCAACAAAGTATGACTTCACTTCAAGTCAACTCAAAACCTTC 646
QY 141 SerAlaAlaGlnMetGlyValAlaGlyValAlaValAlaAlaAlaLeuSerGlyIleSer 160
DB 647 AGTCCGCGCAATGAAAGAGTCGAAGCGGTGTTGTTGCTGCTTCAGGAAAAAGT 706
QY 161 SerGlySerAlaIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 180
DB 707 TCGGTTCCGCAAAATTTGGAACACCTGAGCTCCCAAGCGCGGTGACACCAATATCA 766
QY 181 GluValIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 200
DB 767 GAGTTATCGAAATCGGACTCGCGCTTCTTAAAGCAATTCAGACATGGAGAGAACGCA 826
QY 201 LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnIleu 220
DB 827 AAATCTGCTTATCTACTATGTAATACACAGCAACGACCAACCAAAATTAACATCA 886
QY 221 GlyLeuGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 240
DB 887 GGTCTGAAAGAGCAAGCGATTAATAATCGATTAAGAACGAGAAATACCAAGATGAAG 946
QY 241 AlaAlaGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIleGlnIle 260
DB 947 GCTGCGAAGAGAGATCTTAAGATCTCGAAGCAACATGATCTGCTCAATCTGTGAAG 1006
QY 261 IleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGly 280
DB 1007 ATCGCGGTTCTGTTCCATTCACAGTATTTCTATTGCTGCTGCTATTTTTCATTCGCGGA 1066
QY 281 AlaGlyIleuAlaGlyIleuAlaAlaGlyAlaAlaValaGlyAlaAlaAlaGlyIle 300
DB 1067 GCTGACCTGCTGAGACTCTGCTGGGAGCTGCTGAGTGAAGCGGCGGCTGAGAGTGA 1126
QY 301 AlaGlyAlaAlaAlaAlaAlaThrValAlaAlaThrGlnIleThrValGlnAlaValGln 320
DB 1127 GCAGGAGCTGCTGCCCAACCAACGAGTACCAACCAATTAAGTTCACCTGTTGCCAA 1186
QY 321 AlaValIleGlnAlaAlaAlaIleThrValAlaAlaAlaAlaIleThrAlaAlaIle 340
DB 1187 GCGGTGAACAAAGCTGTTTTCACAGCTGTCAGACAAAGCATCCGCGCTTAAGAGG 1246
QY 341 AlaValIleSerGlyIleIleAlaPheIleIleThrIleValIleAlaIleAlaIle 360


```
CDS
8267..9520
/feature="Cp0814"
/codon_start=1
/transl_table=11
/product="CT573 hypothetical protein"
/protein_id="AAD18952.1"
/db_xref="GI:4377126"
/translation="MAGYSPNYVDLYRISLSAESRLGIRAPQAESLLGCALANF
PDLEERLPDLKRELLFLGSDRGKFLQSLAESKECYIALKREVLNVTSSR
GPVSESPKGVPELIMTECSVSDPAGREVVRLQGLAKELISPRDCELTFLNPN
KLDCWEINGFRVDASEFYKQKIRIRGVKPLLMHGATYAKATKTERDVPSSDENY
SRVLAVGVLLMDGNCWQTCGFQDASSRAPLFEVKRIDDKVMAIDLVNNGVQKQTL
SLVKGVPSPRIINEVIRIEFTGMRSMKSPYLVGGORLILSPDMVLTAKGMKLS
RADIOIDYVTKGVTKGPIILVFEKLEKDLRGFVLRGHMFNAQRITVELISLPKQGEPA
VASQESVNTSRSAAHPGATNRNGS"
9521..11785
/feature="gspe"
/feature="Cp0815"
9521..11785
/feature="gspe"
/codon_start=1
/transl_table=11
/product="General Secretion Protein D"
/protein_id="AAD18953.1"
/db_xref="GI:4377127"
/translation="MVFPRNSLHFLVALSGMLCCSSGVALTIAEKMSLHSGRGAD
YEGASFANMREYSLOLSKLYEARKIRASGTEPKALMDLIRIGVRYLRIEE
LMAEIRKRGKNGLEBYALMNPETTYLNLVDYGEDSYILPQIGIKRATLSKV
YPKSEFEDCLQILSRDLIGYRQVNSMIRKELTMKKECSVAQVSSKDLALPETA
YIGVPLNSVDAPHNQHLKFINETTHVDVIAKRWFGSAGVGBLLKYNVQS
ESIRQEVVILPTKIDPCGEMISILNAAFREDVTKGVSEESGLRVPLQYGRSLFS
GTALVQALFLIRELEGIEINPTKTVFWMVKHSDPOELALISOVHDYFSGENKA
SVGADGCGSOLNASIOIDTYSSAKGSVYGNFIDSKGTILIMVKEKVLPRIO
MLKRLDYPKRWIRVLEILLFEKRLAHBQSGNLRLRGEVCKKCGKSPSSMAGTGI
LELFKSGTSSIVGYDLAYOFLMAQEDVRINASPVTYTNQTPARLAVYDEMIAV
SSDKDAQYNRAQYGIKIMKLPVINVGEDKSYITLETIDFTTGNHNDPRPVT
RNITNKVRIADGEIVIGLRCKMDSHDGIPFLGDIPIGIGIKFGMSTSDSLTEM
VFETRIKENVEQDERKEEALLSRPEREYVQALAASEAAARAHHKLEMPASG
VLSQVEROEYDGC"
11775..13265
/feature="gspe"
/feature="Cp0816"
11775..13265
/feature="gspe"
/codon_start=1
/transl_table=11
/product="General Secretion Protein E"
/protein_id="AAD18954.1"
/db_xref="GI:4377128"
/translation="MAASILSQELDILPYPLKHKCLPIEESSEATIIAHATATSV
IAODEVKLILKKPVRFVLEKESEILQRLQOLVSNREGVNSDKLTMKEEDGTTISEE
DLLETPTDIPVYRLNLWLKAEIERRADIHREPEDSMKRIYRIDVULHDPHSPSH
LRSALTFLKVLAKMDIAEHLRPQGRKRIHGGVEGDMKSTVYITIGERVALLD
KRNVLIDLAGLHMPKGTIELFKDITTAEGILLVGPGSGSTTLVLEKLPILT
NIMTIEDPEPKLPGIAQIAVAPKIGLTFARGLRHLRQDDPILVNGEIRDETEAIA
IOALTGLHVVSTGLTNDALISAIPRLDMGISEYLSLTIIVVAVQRLVRTCPCKCA
ATPENQKSFLLASGKIDTEMPLRYRGQGVHCFRSGYGRGSIYEFLRPNTLFRSEVA
```

Alignment Scores:

Pred. No.:	1.05e-118	Length:	16903
Score:	2369.00	Matches:	496
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	1	Gaps:	0

US-09-889-314-2 (1-496) x AE001662 (1-16903)

QY 1 AsptHrAsmMetSerLleSerSerSerGlyProAspAsnGlnLysAsnLleMetSer 20
DB 3954 GATACAAACATGCTTATCTTCTCAAGACCTGACACATCAAAAATATCATGTCT 3895

QY 21 GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGln 40

```
|||||
Db 3894 CAAGTCTGCATCATCACACACCCAGGCGCGCCCAACAAGATAAAGCTGCTGCACACGAA 3835
QY 41 ThrLysGlnLleGlnGlnThrArgGlnGlnLysAsnThrGluMetGluSerAspLleThr 60
Db 3834 ACGAAGCAAAATACAGCAAAACAGCTCAGGCTAAACACATGAGATGGAAGCGCATCCACT 3775
QY 61 IleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaPro 80
Db 3774 ATTGCTGGTCTCTTGGAAGAAACAAACTTCTGACATCAAAAAACAGAACAGACTGCA 3715
QY 81 GlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyLysAspThr 100
Db 3714 CAACAGGAGGAGTTCGCTCGGAAAGAAATCCTGAGAAAGCAAAAGCAGAGCTGCTACT 3655
QY 101 GlyValSerGlyValAlaAlaThrThrAlaSerAsnThrAlaThrLysLleAlaMetGln 120
Db 3654 GGAGTATCAGAGAGCGGCTCTACTACAGCATCAAAATACGCAAAATTCCTATGCAG 3595
QY 121 ThrSerLleGlnGlnLysSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeu 140
Db 3594 ACCTCATTTGAAGAGCGGAGCAAAAGTATGAGTCTACTTACAGTCACTTCAAAAGCTTC 3535
QY 141 SerAlaAlaGlnMetLysGlnValGlnAlaValValAlaAlaLeuSerGlyLysSer 160
Db 3534 ACGGCGCGCCCAATGAAAGCAAGTCGAAGCGGTGTGTTGCTGCGCTCTCAGAGAAAGCT 3475
QY 161 SerGlySerAlaLysLeuGlnThrProGluLeuProLysProGlyValThrProAspSer 180
Db 3474 TCGGGTTCGCAAAATGTGAACACCTGAGCTCCCAACCGCGGGGTGACCCAAACATCA 3415
QY 181 GlnValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlyAlaThr 200
Db 3414 GAGGTATATGAAATCGAATCGACTCCGCTGCTTAAAGCAATTCAGACATTTGGAGAACGCA 3355
QY 201 LysSerAlaLeuSerAsnTyraLleSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
Db 3354 AATCTGCTTATCTAATATGCAAGTATACCAAGCACAAGCAGACCAACAAATAACTCA 3295
QY 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGluArgGlnGlnLysThrGlnLysMet 240
Db 3294 GGTCTGAAAGCAAGCAAGCAATGAAAGCAATGAAAGCAAGCAAGCAAGCAAGCAAG 3235
QY 241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMet 260
Db 3234 GCTGCCGAACAGAGCTTAAAGTCTCGAAGACATGATGATCTCATATCTGATGATG 3175
QY 261 IleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIleIleThrCysGly 280
Db 3174 ATCGCGGTTTCTGTTCCATTCACAGTTATTTCTATTGTTGCTGCTATTTTACATCGCA 3115
QY 281 AlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaAlaGlyAla 300
Db 3114 GCTGGACTGCTGCGAGTCTCCCTCGGAGGCTGCTGAGGTCGACGCGGCTGAGGTGCA 3055
QY 301 AlaGlyAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGln 320
Db 3054 GCAAGAGCTGCTGCGCAACCAAGCAAGCAAGCAAAATTCACCTTAACCTGTTGCTCAA 2995
QY 321 AlaValLysGlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAla 340
Db 2994 GCGGTGAACCAAGCTGTATTCACAGCTGTACAGCAAGCATCACCGCGCTATTAAGAAGCG 2935
QY 341 AlaValLysSerGlyLysAlaIleLysAlaIleLysThrLeuValLysAlaIleAlaLysAla 360
Db 2934 GCTGTCAAAATCTGGAAATTAAGCAATTAATCAAAACTTGTAGTCAAAACGATGTGCAAAAGCC 2875
QY 361 IleSerLysGlyLysSerLysValPheAlaLysGlyThrGlnMetLleAlaLysAsnPhe 380
Db 2874 ATTTCTAAAGGAATCTCTAAGGTTTTCGCTAAGGAACTCAAAATGTTGCGAAGAACTTC 2815
QY 381 ProLysLeuSerLysValIleSerSerLeuThrSerLysTryValThrValGlyValGly 400
|||||
```

[illegible]

AX033283
LOCUS AX033283 1491 bp DNA linear PAT 21-SEP-2000
DEFINITION Sequence 1 from Patent WO0046359.
ACCESSION AX033283
VERSION AX033283.1 GI:10280098
KEYWORDS
SOURCE Chlamydomonadales.
ORGANISM Chlamydomonadales; Chlamydiales; Chlamydiales.
REFERENCE
AUTHORS Matthews, R.C. and Burnie, J.P.
TITLE Chlamydomonadales; Chlamydiales; Chlamydiales.
JOURNAL
MATTHERS RUTH CHRISTINE (GB) ; NEUTECH PHARMA PLC (GB) ; BURNIE JAMES P (GB)
FEATURES
source
location/Qualifiers
1..1491
/organism="Chlamydomonadales pneumoniae"
/db_xref="taxon:83558"
/note="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CAC09981.1"
/db_xref="GI:10280098"
/translation="DTNMSISSSGPDNOKNIMSOVLTSTPOGVPOODKLSNETKOI
OOROGKTEMSDATTAGASGDKXTSTTETETAPOGVAAKESSESOKAGADTV
SGAAATTSNTATRIAMOTSTIEBASKSHESTLESLOSIAAOKMEVEAVVAALSGS
SGSAKLEIPELPFGVTPRSERIEITGLALAIOTLGATTSALSNVSTOQADOTN
KLGLEKQAIKIDKREVEYOEKMAEOKSKDEGMTDVTYVAVNAIVSIVAI
FTCGAGLAGLAGAAGVAAAGAAAGAAATVATQIVQAVVOVKQAVTAVQAI
TAAIKAAVKSIGKAFIKTLVKAIAKAIKSKISKVTOMIAKIPPLSKVLSLTS
KMTVGVGVVVAAPALGKIGIMOMOLEMOQNVADOKRFGKLAADAMISMTQEMQO
ASKTASKOTGSGSNEMPOKATKLGAIILKAVAIISALINAKHTNPF"

BASE COUNT 485 a 335 c 348 g 323 t
ORIGIN
Alignment Scores:
Pred. No.: 5,64e-120 Length: 1491
Score: 2369.00 Matches: 496
Percent Similarity: 100.00% Conserved: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: Gaps: 0
US-09-889-314-2 (1-496) x AX033283 (1-1491)

QY 1 AspThrAsnMetSerIleSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20
DB 1 GATACAAACATGCTATTCATCTCTTCAGAGACCTGACAAATCAAAAATAATCATGTCT 60
QY 21 GlnValIleuThrSerThrProGlnGlnValProGlnGlnAspLysLeuSerGlnGlu 40
DB 61 CAAGTTCGACATTCACACCCAGGGCGTCCCAACAAAGATTAAGCTGCTCGGCAACGAA 120
QY 41 ThrLysGlnIleGlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThr 60
DB 121 ACGAAGCAATATACGCAACACGCTGAGGTAAACACATGCAATGCAAGCCATGCTCACT 180
QY 61 IleAlaGlnLysSerGlyLysAspLysThrSerSerThrThrLysThrGlnLysAlaPro 80
DB 181 ATTGCTGGTCTTGGAAGAACAAACTTCCTCGACTACAAAACAGAACAGCTCA 240
QY 81 GlnGlnGlnValAlaAlaGlnLysGlnSerGlnSerGlnLysAlaGlnLysAlaAspThr 100
DB 241 CAACAGGGAGTGTGCTGGGAAAGATCTCAGAAAGTCAAAAGGACAGCTCTGACTACT 300
QY 101 GlnLysSerGlnAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120
DB 301 CGAGTATTCAGGAGCGGCTCTCTACAGCATCAATATCTGCAACAAATTTGCTATGCG 360
QY 121 ThrSerIleGlnGlnLysSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeu 140
DB 141

DB 361 ACCCTATTGAAGAGCGACGCAAAAGTATGAGCTACCTTAGACTTCAAAAGCCTC 420
QY 141 SerAlaIleGlnMetLysGlnValGlnValAlaValAlaIleAlaLeuSerGlyLysSer 160
DB 421 AGTGGCGGCGCAATATGAAAGTCAAGGCGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
QY 161 SerGlnSerAlaLysLeuGlnThrProGlnLysProLysProGlnValThrProLysSer 180
DB 481 TCGGCTCCGCAAAATTTGAAACACCTGAGCTCCCAAGACCGGGGTGACACCAAGATCA 540
QY 181 GlnValIleGlnIleGlnLysLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlnAlaThr 200
DB 541 GAGTTATATGAAATGAGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 600
QY 201 LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
DB 601 AATCTGCTTATCTTAACATCTCAAGTACACAGACCAAGACCAACCAATTAACCA 660
QY 221 GlnLeuGlnLysGlnAlaIleLysIleAspLysGlnLysGlnGlnGlnGlnGlnGlnGln 240
DB 661 GGTCTAGAAACCAAGCCTTAATAATCGATTAAGAACGAAAGCAATACCAAGATGAAG 720
QY 241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlnLysMetAspThrValAsnThrValMet 260
DB 721 GCTGCCAAGCAGAGCTTAAGATCTCGAAGGACCAATGATGATCAATGATGATGATGATG 780
QY 261 IleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIleIleThrGlyGly 280
DB 781 ATCGGGGTTCTGTTGCCATTCATCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 840
QY 281 AlaGlnLysLeuAlaGlnLysLeuAlaGlnLysAlaValAlaGlnAlaAlaAlaGlnGln 300
DB 841 GCTGAGCTGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
QY 301 AlaGlnAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValGln 320
DB 901 GCAGAGAGCTGCTGCCAAGCAGGACGACGACCAACCAATTAAGCTGCTGCTGCTGCTGCT 960
QY 321 AlaValLysGlnAlaValIleThrAlaValAlaGlnAlaIleThrAlaAlaIleLysAla 340
DB 961 GCGGTGAACCAAGCTGTTATCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
QY 341 AlaValLysSerGlyIleLysAlaIleLysAlaIleLysAlaIleLysAlaIleLysAla 360
DB 1021 GCTGCTCAATCTGGAATTAACATTAACATTAACATTAACATTAACATTAACATTAAC 1080
QY 361 IleSerLysGlyIleSerLysValIleAlaLysGlnLysGlnMetIleAlaLysAsnIle 380
DB 1081 ATTCTTAAGGAATCTCTAAGCTTTCCTTAAGGACCAATGATTCGCAAGAACTTC 1140
QY 381 ProLysLeuSerLysValIleSerSerLeuThrSerLysThrValIleGlnValGln 400
DB 1141 CCCAAGCTCTCGAAGATCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
QY 401 ValAlaValAlaAlaProAlaLeuGlnLysGlyIleMetGlnMetGlnLeuSerGlnMet 420
DB 1201 GTTGAAGTGGCGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
QY 421 GlnGlnAsnValAlaGlnIleGlnLysGlnValGlnLysLeuGlnAlaAlaAlaAspMet 440
DB 1261 CAACAAAGCTGCTCAATTTGCAAGAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 1320
QY 441 IleSerMetPheThrGlnPheThrPheGlnGlnAlaSerLysIleAlaSerLysGlnThrGly 460
DB 1321 ATTCTATGTTCACTCAATTTGCAAGAGGCAAGGCAAGGCAAGGCAAGGCAAGGCAAGG 1380
QY 461 GluSerAsnGlnMetThrGlnLysAlaThrLysLeuGlnAlaGlnIleLeuLysAlaThr 480
DB 1381 GAGTCTAATGAATATGACTAAAGACCTACCAAGCTGCGGCTCAATCTTAAACGCTAT 1440
QY 481 AlaAlaIleSerGlyAlaIleAlaGlnAlaLysLysThrAsnAsnIle 496
DB 1441 GCCCAATATGAGGAGCCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1488

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 27, 2003, 14:06:32 : Search time 2670 Seconds

(without alignments)
5406.364 Million cell updates/sec

Title: US-09-889-314-2
Perfect score: 2369
Sequence: 1 DTNMSISSSGPDNOKNIMS.....LKAYAAISGAIAGAHKTNMF 496

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2054640 segs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL-frame+pn.model -DEV-rlh
-Q/cgn2.1/USPTO.spool/US09889314/runat.24012003.144540.10107/app.query.fasta.1.647
-DB-GenEmbl -QFMT-fastlap -SUFFIX-rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS-bits -START=1 -END=1 -MATRIX-biosum62 -TRANS-human40.cdi -LIST=45
-DOCAIIGN=200 -THR.SCORE-pct -THR.MAX=100 -THR.MIN=0 -ALIGN=45 -MODE-LOCAL
-OUTFMT=pro -NORR-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER-US09889314.eccn.1.1.1616@runat.24012003.144540.10107 -NCPU=6 -ICPU=3
-NO_XLUPY -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -LONGLOG -DEV_TIMOUT=120
-WARN_TIMOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELop=6 -DELEXT=7

Database : GenEmbl.*
1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vl.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.ro.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg_hum.*
31: em.htg_inv.*
32: em.htg_other.*
33: em.htg_mus.*
34: em.htg_pin.*
35: em.htg_rod.*
36: em.htg_mam.*
37: em.htg_vrt.*
38: em.sy.*
39: em.htgo_hum.*
40: em.htgo_mus.*
41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2369	100.0	1491	6	AX033283	AX033283 Sequence
2	2369	100.0	16903	1	AE001662	AE001662 Chlamydia
3	2356	99.5	1939	6	E16639	E16639 Chlamydia p
4	2356	99.5	1939	6	E16674	E16674 gDNA encodi
5	2356	99.5	15455	1	AE002263	AE002263 Chlamydo
6	2340	98.8	325865	1	AP002548	AP002548 Chlamydo
7	2326	98.2	9595	6	E16678	E16678 All sequenc
8	2320	97.9	11099	6	E16636	E16636 Expression
9	2317	97.8	1587	6	E16677	E16677 All sequenc
10	2316	97.8	1464	6	E16671	E16671 DNA encodin
11	2316	97.8	1464	6	AR122105	AR122105 Sequence
12	2316	97.8	1464	6	E12520	E12520 DNA encodin
13	2316	97.8	1464	6	E12535	E12535 Chlamydia p
14	2316	97.8	1464	6	E12541	E12541 DNA encodin
15	2316	97.8	1464	6	E16635	E16635 Chlamydia p
16	2316	97.8	1947	6	AR122114	AR122114 Sequence
17	2316	97.8	1947	6	E12525	E12525 DNA encodin
18	2316	97.8	1947	6	E12538	E12538 DNA encodin
19	1282	54.1	5658	6	AR122110	AR122110 Sequence
20	1282	54.1	5658	6	E12523	E12523 Nucleotide
21	1282	54.1	5658	6	E12544	E12544 Nucleotide
22	1265	53.4	5438	6	AR122122	AR122122 Sequence
23	1265	53.4	5438	6	E12527	E12527 Nucleotide
24	1265	53.4	5438	6	E12540	E12540 Nucleotide
25	1262	53.3	1048	6	AR122109	AR122109 Sequence
26	1262	53.3	1048	6	E11479	E11479 DNA encodin
27	1262	53.3	1048	6	E12265	E12265 DNA encodin
28	1245	52.6	777	6	AR122107	AR122107 Sequence
29	1245	52.6	777	6	E12522	E12522 DNA encodin
30	1245	52.6	777	6	E12543	E12543 DNA encodin
31	1245	52.6	813	6	AR122106	AR122106 Sequence
32	1245	52.6	813	6	E12521	E12521 DNA encodin
33	1245	52.6	813	6	E12542	E12542 DNA encodin
34	1245	52.6	1296	6	AR122115	AR122115 Sequence
35	1245	52.6	1296	6	E12526	E12526 DNA encodin
36	1245	52.6	1296	6	E12539	E12539 DNA encodin
37	1233.5	51.6	9983	1	AE001328	AE001328 Chlamydia
38	1223.5	51.6	10653	1	AE002353	AE002353 Chlamydia
39	238.5	10.1	179553	2	AC024253	AC024253 Homo sapi
40	220.5	9.3	5393	6	ST51PGNS	ST51PGNS S. typhi sp
41	220.5	9.3	5393	6	AR082442	AR082442 Sequence
42	220.5	9.3	274050	1	AL627276	AL627276 Salmonella
43	218.5	9.2	2995	1	SDU66877	SDU66877 Salmonella
44	218.5	9.2	3622	1	STU25631	STU25631 Salmonella
45	218.5	9.2	3622	6	AR082441	AR082441 Sequence

RESULT 1

ALIGNMENTS


```

Db 18606 AAACAGATATATAGTTAAAGATTAGAGTGCAGAAC-----ATGAGGTGATGATCAGTA 18659
      |||          :|||:|||||
      61 lIealagIyAlaSerGlyLysAspLysThrSerSerThrThrlLysThgluThralapo 80
      :|||          :|||:|||||
Db 18660 GTAATTGAGATTAGTTTACATGATTAATAAAGCAGAT-----AAACACTGGATCGTTT 18713
      :|||          :|||:|||||
      81 GInglngIyAlaIalagIyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThr 100
      :|||:|||||
Db 18714 GACACAGATTTAGCG-----AAAGCAGCGCAAAAT--- 18743
      :|||:|||||
      101 GYAlaSerGlyAlaAlaIalathrThralaSerAsnThrAla---ThrlYsIlleAlaMet 119
      :|||:|||||
Db 18744 -----GCAGGGCGCGCATTAATAGCGTATAGAGAGCGGTATTCGATTTTCGTAAGT 18797
      :|||:|||||
      120 GlnThr-----SerIleGluIuAlaSerLysSerMetGluSerThrIleu 134
      :|||:|||||
Db 18798 CATCGAAGCGATTAAAGACGCTTGTAAATGCGTTTAAAGCATGGGAGTGGCTGGC 18857
      :|||:|||||
      135 GUSerIleuGlnSerLeuSerAlaIalagImetLysGluValGluAlaValAlaIa 154
      :|||:|||||
Db 18888 TCAAATGCTTTAAAGCTATGTTAACTTTATGCGTGAATGCTTCAATGATCAGAGCC 18917
      :|||:|||||
      155 AlaleuSerGlyLysSerSerGlySerAlaLysLeuGluThrProGluLeuProLysPro 174
      :|||:|||||
Db 18918 GCTTATCTAAACTGATCA----- 18938
      :|||:|||||
      175 GYValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGln 194
      :|||:|||||
Db 18939 -----ACAGTAAAGACTGGTTCGTAAACGCTGCTAAACATCTATATACAGCAAAAG 18992
      :|||:|||||
      195 ThrLeuGlyGluAlaThrLysSerAlaLeuSerAsnTyAlaSer-----ThrlGln 211
      :|||:|||||
Db 18993 GAACCTGGAAACAAGTATCAAAACACAGCGGTTAATATTAAACGCGCTTCTTCAATT 19052
      :|||:|||||
      212 AlaglnAlaAspGlnThrAsnLysLeuGlyLeuGluLysGlnAlaIleLysIleAspLys 231
      :|||:|||||
Db 19053 GCTAAGACAGCTCAAAAGTAGTATTGTGTACGCTTAAAGTTCATATATGTCATTA 19112
      :|||:|||||
      232 GluArgGluGluThrGlnGluMetLysAlaIalagIuGlnLysSerLysAspLeuGluGly 251
      :|||:|||||
Db 19113 TCCATCCCGCGCAATGTAAAGCGCTGGAATCAGTATTAAATCAACA---TTAGTAACT 19169
      :|||:|||||
      252 ThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSer 271
      :|||:|||||
Db 19170 AGTTTGCAGACGCTAAATCGGCT-----GCTATTCTTTTGTCTCAACTGTA 19223
      :|||:|||||
      272 lIeValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaIalagIyAlaIa 291
      :|||:|||||
Db 19224 GTTATTAAAGTATT-----CCAGGAGCTGCT 19250
      :|||:|||||
      292 ValGlyAlaIalAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 309
      :|||:|||||
Db 19251 AAACAGCGCGCTACAGCAGTAAACAGTTCGTAGTACCAAAAGCGGTGTAGT 19310
      :|||:|||||
      310 AlAThGlnIleThrValGlnAlaIaValAlaIaValLysGln----- 324
      :|||:|||||
Db 19311 GCTGCTTATATGAGCGCTTAAGAGAACTATTAGCGCTGTGAAGCTATTCTTACGCTACA 19370
      :|||:|||||
      325 -----AlaValIleThrAlaValAlaArgGlnAlaIleThrAlaIalAlaIleLys 339
      :|||:|||||
Db 19371 AATATCGGCGCATTTAGCATTAAGTACGAGCAATGAAGAAAGCTTTTACGCTGATCAAGC 19430
      :|||:|||||
      340 AlAlaIaValLysSerGly-----lIeLysAlaPheIleLysThr-----LeuVal 354
      :|||:|||||
Db 19431 GCGCGCAAAACACAGCAAGCAACAGTGAATACAGTATTAACACGCTTACCGCTGG 19490
      :|||:|||||
      355 LysAlaIleAlaIaLysAlaIleSerLys---GlyIleSerLysValPheAlaLysGlyThr 373
      :|||:|||||
Db 19491 AATATCGGCGCAAGCGCGCAGCCAAAGCTGGATTTCAGCATTA-----AAAGCGCTTA 19544
      :|||:|||||
      374 GlnMetIleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLys 393
      :|||:|||||

```

```

Db 19545 GGAACATTGCGAAGAACACGCTGTTCTTAATTAA-----AGTGATTAAGTAACGGA 19598
      :|||:|||||
      394 TrpValThrValGlyValGlyValAlaIalAlaProAlaLeuGlyLysGlyIleMet 413
      :|||:|||||
Db 19599 TTT-----AACGCGCAAAAGCGCGGAAAGTGACAGGCGCTGGAAATGCGT 19646
      :|||:|||||
      414 GlnMetGlnLeuSerGlnMetGlnGlnAsnValAlaGlnPheGlnLysGluValGlyLys 433
      :|||:|||||
Db 19647 GAAGCACTTAAATTCAGTTTCAAAAGCCCGCGCAACAAAGCTCGCTTATGATTCTCAGA 19706
      :|||:|||||
      434 LeuGlnAlaAlaAlaAspMetIleSerMetPhe----- 444
      :|||:|||||
Db 19707 TTACGACGACGCTGGATTAATTCACAGCACTAAAAATGTTGGTAGCCCTATTTGCT 19766
      :|||:|||||
      445 ---ThrlGlnPheTrpGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyLysSerAsn 463
      :|||:|||||
Db 19767 CGAGTTGATACGATTGATACCTCAACTAAATGTTAACAGTCTTACGCTTACGCTACCAAAA 19826
      :|||:|||||
      464 GluMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaIaIaIaIaIaIa 483
      :|||:|||||
Db 19827 GATGCGCAACTGATTATGACACACCTTACAGCCGCTATCGATGCGACACCATCCGATTA 19886
      :|||:|||||
      484 SerGlyAlaIleAlaGlyAlaHisLys 492
      :|||:|||||
Db 19887 GATGCTGTGCGCATTAGGCTTAAAAA 19913
      :|||:|||||
RESULT 44
AB069245
ID AB069245 standard; DNA; 3011208 bp.
XX
XX
XX AB069245;
AC
XX
XX 29-AUG-2002 (first entry)
DT
XX
XX Listeria innocua DNA sequence #64.
DE
XX
XX Antibacterial; Listeria; food contamination; mutational analysis;
KW
XX infection; ds.
XX
XX Listeria innocua.
OS
XX
XX WO200228891-A2.
PN
XX
XX 11-APR-2002.
PD
XX
XX 04-OCT-2001; 2001WO-FR03061.
PF
XX
XX 04-OCT-2000; 2000FR-0012697.
PR
XX
XX (INSP ) INST PASTEUR.
PA
XX (CNRS ) CNRS CENT NAT RECH SCI.
PI
XX
XX Kunst F, Glaser P;
XX
XX WPI: 2002-332479/37.
XX
XX New genomic sequences from Listeria species, useful for detection,
PT treatment and prevention of infection, also related polypeptides,
PT antibodies and modulators
XX
XX Claim 5; SEQ ID 2058; 180pp; French.
XX
XX The present invention relates to nucleic acid sequences
CC (AB067188-AB071212) from Listeria sp. The sequences are useful as probes
CC and primers for identification and/or detection of Listeria (e.g. as
CC contaminants in foods, or mutational analysis) and for analysis of
CC gene expression. Proteins encoded by the nucleic acid sequences can be
CC used to screen for compounds that modulate gene expression, replication
CC and pathogenicity of Listeria (potential therapeutic agents), also for
CC treating infections by Listeria, and are useful as immunogens in
CC anti-Listeria vaccines.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format

```



```

Db 1807 TCCACAGCGCTGATTATTAAATCCGACAGCTGCCCTCAGGT-----GGAGCAACGCTGTCA 1754
Oy 109 ThrAlaSerAnThrAlaThrLysIleAlaMetGlnThrSerIleGlu---GluIaSer 127
Db 1753 ACTGCAACTTCTGTGTACTTCCGACAGCGGTGCAACTGGCGTTTCAATGGCAGCTTCT 1694
Oy 128 LysSerMetGlnSerThrLeuGlnSerLeu---GlnSerLeuSerAlaAlaGlnMetLys 146
Db 1693 GATAGTACAGAAACATCATAGTATGCCGACACTGCCGAGGTGGTGGTGTCTCG 1634
Oy 147 GluValGluAlaValAla-----ValAlaAlaLeuSerGlyLysSer 160
Db 1633 GGAACAGAGCGCTGTGTGTGGAGTAGAAACGAGGAGCGGACGCGACGACCTTCA 1574
Oy 161 SerGlySerAlaLysLeuGlnThrProGluLeuProLysProGlyValThrProArgSer 180
Db 1573 GCGGTGCATCCAGAG-----GCTCG 1553
Oy 181 GluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThr 200
Db 1552 GCGATCACAGGACCACTTGTGTATCTAAAGCTGCTATGCTCAGTGGTCTGATGCT 1493
Oy 201 LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
Db 1492 GGTGCTGTCTCTCAGAGTTCTTCAGCCACTACAGCTGTGTACAGTTCGCGACACCA 1433
Oy 221 Gly-LeuGluLysGlnAlaIleLysIleAspLysGluArgGluLysGlnGlnMet-- 239
Db 1432 GGAGGAGAACAGCTGGCGCTA-----AGTCACAGGAGGTTTTCACACAGCA 1385
Oy 240 -----LysAlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAs 257
Db 1384 CTAGAACAGCGGTGATGAGCGATCACCGAGGTATTCG--GGAACAGCAGATGAGTGAC 1326
Oy 257 nThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaIlePhe 277
Db 1325 ATCCGATCCGTTACTGCTGATGTCAGGAAGT-----GGTGTGCTGGC 1281
Oy 277 eThrcysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaValAlaGly-----Al 294
Db 1280 AACGACTGGCGGTGATGATTCGCGCCACAGAGGTACACAGGGGTGGTGTCTC 1221
Oy 294 AlaAlaAlaGlyAlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 314
Db 1220 AGCAACAGCACTGAAGCAACTTTCGGGTTTCTGCACACAGGTCGCAATGGTTTC 1161
Oy 314 rValAln---AlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValAlaArgGlnAl 333
Db 1160 TGGCTCTGGGCTGAGAGAGGTAGCTAACACAGTATTCGCAACTTGGGACTC 1101
Oy 333 AlIleThrAlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLe 353
Db 1100 GGCACAGACTGAGCAGAGGCGTGGCTT---TCTGTGACAGTTGTAGAGTGCACAGACC 1044
Oy 353 uValLysAlaIleAlaLysAlaIleSerLysGlyLysSerValPheAlaLysGlyTh 373
Db 1043 TACCGGGGCTGAGTGAAGCACTTACAGAGCAGCAGCACTGGGCTCCACAGCTGTGC 984
Oy 373 rGlnMetIleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLys 393
Db 983 A-----ACCGACAC 975
Oy 393 sThrP-----ValThrValGlyValGlyValAlaValAlaAlaProAl 407
Db 974 ATGGGGGAGAGCTGCAGCCGAGCAACAGCTGGGTTTCAGTAGCCGGTGGAGCAACGAC 915
Oy 407 AlenGlyLysGlyIle-MetGlnMetGlnLeu----- 417
Db 914 CACTGGCTACAGTGCAGTAGTGTTCCTGCACAGCAGAAACCTCGGGATGAGAG 855
Oy 418 -----SerGluMetGlnGlnAsnValAlaGlnPheGln- 428
Db 854 TGCACAGCACTCCGGTATGTGCAACCGAATTCACAGACAGGTGTGTGAGCTGTGAGT 795

```

```

Oy 429 -----LysGluValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnP 447
Db 794 CTCGGCAGCAGACGGGGGCACTGCGAGAGCAACGGCAGCGCCAGAACTGGGGCAACAC 735
Oy 447 heTrpGlnGln-----A 451
Db 734 AGAAGACACACCTGTGTACAGCAGTACAGTGTCTGCAACGGGAACCGTGTGTAGCGG 675
Oy 451 lAserLysIleAlaSerLysGlnThrGlyLysSerAsnGlnMetThrGlnLysAlaThrL 471
Db 674 AACCAAGCTGTGTGTGGATC-ACAGGCGAAACACAGCGGAGTACG-----GACGCG 622
Oy 471 yLsLeuGlyAlaGlnIleLeuLysAlaTyrAlaAlaIleSerGlyAlaIleAlaGlnAla 491
Db 621 GAGTGGGGGACAGAGGACACACAGTGTGTGTGTATATCAGAGCGCAGCTACAGGAATTG 562
Oy 491 lLysThr 493
Db 561 GCGCAACA 554

RESULT 42
AAV52503
ID AAV52503 standard; DNA; 750 BP.
AC AAV52503:
XX 23-OCT-1998 (first entry)
DE Streptococcus pneumoniae genome fragment SEQ ID NO:370.
XX Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
KW computer readable medium; vaccine; pharmaceutical composition; ds.
OS Streptococcus pneumoniae.
XX WO9818931-A2.
XX 07-MAY-1998.
XX 30-OCT-1997; 97WO-US19588.
XX 31-OCT-1996; 96US-0029960.
XX (HUMA-) HUMAN GENOME SCI INC.
XX Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;
XX Kunsch CA, Rosen CA;
XX WPI; 1998-272225/24.
XX
XX Computer-readable medium with recorded Streptococcus pneumoniae
XX polynucleotide sequences - useful in diagnostic kits and assays, and
XX pharmaceutical compositions and vaccines for Streptococcus
XX pneumoniae
XX
XX Claim 1; Page 1384-1385; 1409pp; English.
XX
XX The present invention describes a computer readable medium which has
XX the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)
XX recorded on it, or a representative fragment or a sequence at least 95%
XX identical to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in
XX SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from
XX Streptococcus pneumoniae. The present invention also describes an
XX isolated nucleic acid molecule encoding a homologue of any of the
XX fragments of the S.pneumoniae genome (SEQ ID NO:1 to 391) where the
XX nucleic acid molecule is produced by a process comprising: (a) screening
XX a genomic DNA library using as a probe a target sequence defined by any
XX of the sequences in SEQ ID NO:1 to 391, identifying members of the
XX library which contain sequences that hybridise to the target sequence and
XX isolating the nucleic acid molecules from the members; or (b) isolating
XX mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid
XX molecules whose nucleotide sequence is homologous to amplification

```


Qy	112	AsnThrAlaThrIysIleAlaMetGlnThrSerIleGluGlnAlaSer-----LysSer	129
Db	6695	AAATTTGCAATTAAGGTA-----GAAACAGATATCGATCATAGACTTTCACACTCTTCT	6548
Qy	130	MetGlnSerThrIeuGlnSerIeuGlnSerIeuSerAlaAlaGlnMetLysGluValGlu	149
Db	6949	ACGTGATTAATGTTAAAGAGCTTATTAATAATTTGCAATTCCTTGACAGCGGAGGAGA	7008
Qy	150	AlaValValAlaAlaAlaLeuSerGly-----LysSerSerGly	162
Db	7009	GCCCGACGCTTTCACACAGATTACCGACAGTGTTCTGTTAACACTATAATATGTTCTGTG	7068
Qy	163	SerAlaLysLeuGln-----ThrProGluLeuProLysProGlyValThr	177
Db	7069	ATAGCTCGAGTTACCAATACATACCTGATTTTGACTTCCGTCACGAAAGAAAGTAATGTACG	7128
Qy	178	ProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGly	197
Db	7129	GCAAAGAACGAAAA-----AATATTAGCAAAACAGACGAAATATCGAAGATGGA	7179
Qy	198	GluAlaThrIysSerAla-----LeuSerAsnTyrlaSerThrGlnAlaGlnAla	214
Db	7180	GCAGACGCAATCGGAGCCAAATGCTTGTTGAATTAATTTTGGACAC-----GCTGTAGA	7233
Qy	215	AspGlnThrAsnLysLeuGlyLeuGlnLysGlnAlaIleLysIleAspLysGluArgL	234
Db	7234	GATGAAAAAATTCGTGAAGAAAAAGAACAGAAAGTTTAAAACTTTAGACGAAGTTAAC	7293
Qy	235	GluTyrlGlnGluMetLysAlaAlaGluGlnLysSerLysAspLeu-----	249
Db	7294	AAAGACACAGATTAATAAAGTAATATGATGCTACGAAAAAATTTTCAATCAGACAGATT	7353
Qy	250	-----GluGlyThrMet	253
Db	7354	TCTACAGAAAGTACTTCTGTAAAAACCGGATAGAGGAGTACTCAGGAGAAAGATTAAA	7413
Qy	254	AspThrValAsnThrVal---MetIleAlaValSerValAlaIleThrValIleSerIle	272
Db	7414	GCCATTGTGAAGACTTCTGATATTATTGTGAAAAATGTAGATTATACACAGAGACAG	7473
Qy	273	ValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaVal	292
Db	7474	AATATATCACTTCTACTGCGGTGGAGAACGACGACGCTTGCTCC-----	7521
Qy	293	GlyAlaAlaAlaAlaGlyAlaAlaGlyAlaAlaAlaAlaThrThrValAlaThrGln	312
Db	7522	-----GCATCGAGAACAGTGGCGATTACAAATTAATAAAGAAAT	7560
Qy	313	IleThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValArgGln	332
Db	7561	TCCGAGAGTTACTGTGAAATTTCTTTGTGAAGACGCTGAAAAAGTAATGTTAGATCG	7620
Qy	333	AlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThr	352
Db	7621	GATATTACAGAAATGTGCTTTAACGATATCAAGAGTCTGTGAGACATTTG---GGA	7677
Qy	353	LeuValLysAlaIleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGly	372
Db	7678	ATAGAGAGCTGCCCTATGCGAATTAATTTCAATGGAAGATCAATATATCATGTTAAAT	7737
Qy	373	ThrGlnMetIleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSer	392
Db	7738	TCTTAGCTTAATGAAAAAATATTGATGTTATTGTAAAGATTAATCGGAATTTGAGACGC	7797
Qy	393	LysTrpValThrValGlyValGlyValValAlaAlaAlaAlaAlaLeuGlyLysGlyIle	412
Db	7798	GAAGCAAAAGATTAAACCGTGGAGCGGTAGCTCCGAGCAATTATTCACAAACCAAG	7857
Qy	413	MetGlnMetGlnLeuSerGlnMetGlnAsnValAlaGlnIleGlnLysGlnValGly	432
Db	7858	AATGAATTAAT---TCAAGAGTTGAATTTGAGAAAGATTTTTCAAATGAAGA-----	7908
Qy	433	LysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPheThrPcIlnAlaSer	452

Db 7909 -----AAT 7911
 Oy 453 Lys1leAlaserLySGlnThGlyGluSerxnglUmethTrnglnLyAlaThrLySleu 472
 Db 7912 AGAGTACTAGCCCTCTTAAGGAATTGCAAGAAATCAATGTCAAAGTGAAGAAAAGAA 7971
 Oy 473 G1yAlaGln1leuLeuLyAlaTyrAla1a1leSerG1yAla1leAlG1yAla 490
 Db 7972 AACGAGTGACTGCTGAATCTCAAGAGCTTCTGTAGAGCAGTACGACGAGGCA 8025
 RESULT 39
 ID AAV52396 standard; DNA: 3744 BP.
 AC AAV52396;
 XX AAV52396;
 DT 23-OCT-1998 (first entry)
 XX
 DE Streptococcus pneumoniae genome fragment SEQ ID NO:263.
 XX
 KW Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
 KM computer readable medium; vaccine; pharmaceutical composition; ds.
 XX
 OS Streptococcus pneumoniae.
 XX
 PN MO9818931-AZ.
 PD 07-MAY-1998.
 XX
 PF 30-OCT-1997; 97WO-US19588.
 XX
 PR 31-OCT-1996; 96US-0029960.
 XX
 PA (HUMAN) HUMAN GENOME SCT INC.
 XX
 PI Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;
 PI Kunsch CA, Rosen CA.
 DR WPI: 1998-272225/24.
 XX
 PT Computer-readable medium with recorded Streptococcus pneumoniae
 PT polynucleotide sequences - useful in diagnostic kits and assays, and
 PT pharmaceutical compositions and vaccines for Streptococcus
 PT pneumoniae
 XX
 PS Claim 1: Page 1289-1291; 1409pp: English.
 XX
 The present invention describes a computer readable medium which has
 CC the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)
 CC recorded on it, or a representative fragment or a sequence at least 95%
 CC identical to SEQ ID NO:1 to 391. The nucleotide sequences depicted in
 CC SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from
 CC Streptococcus pneumoniae. The present invention also describes an
 CC isolated nucleic acid molecule encoding a homologue of any of the
 CC fragments of the S.pneumoniae genome (SEQ ID NO:1 to 391) where the
 CC nucleic acid molecule is produced by a process comprising: (a) screening
 CC a genomic DNA library using as a probe a target sequence defined by any
 CC of the sequences in SEQ ID NO:1 to 391, identifying members of the
 CC library which contain sequences that hybridise to the target sequence and
 CC isolating the nucleic acid molecules from the members; or (b) isolating
 CC rRNA, DNA or cDNA produced from an organism, amplifying nucleic acid
 CC molecules whose nucleotide sequence is homologous to amplification
 CC primers derived from the fragment of the S. pneumoniae genome to prime
 CC the amplification and isolating the amplified sequences. The computer
 CC readable medium can be used in a computer-based system for identifying
 CC fragments of the S. pneumoniae genome of commercial importance, or
 CC expression modulating fragments of the S. pneumoniae genome. Products
 CC from the present invention can be used in diagnosis kits and assays, and
 CC pharmaceutical compositions and vaccines for S. pneumoniae.
 XX
 Sequence 3744 BP; 1119 A; 804 C; 773 G; 1048 T; 0 other;

```

Db 3880 TCTACAGCTGCTCAACATCATTAAGTAGATTCACAGAGTGTAGCAAGTCCAAAT----- 3933
Oy 209 SerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlyLysGln----- 225
Db 3934 TCACTAAGTGCATCATATGTCGACCAAGTGTATCATCATCAAGAAATCAGAGTCATTA 3993
Oy 226 -----AlaIleLysIleAspLysGluArgGlyIuArgGlyIuArgGlyIuMetLysAla 241
Db 3994 AGTCATGAGCATCATTAAGCGTTCAACAGAGTAGAGTAGAATCGCGGCACATCATCA 4053
Oy 242 AlaGlyGlnLys-----SerLysAspLeuGlyIuThrMetAspThrValAsn 257
Db 4054 AGTGAAGTAAATCCGATTCAACATGAGCCCTGACATGCTCCCAATCAACAGCAGA 4113
Oy 258 ThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePhe 277
Db 4114 AGTACAGTGTGTCAACATCAAGAAAGTTGTCTGACTCAACAGATCAATCATGTCAC 4173
Oy 278 ThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValAlaAlaAlaAla 297
Db 4174 AGTGCCTCAATGATCAACAAAGTGTGTAGACTCAAACTCAGCAAGCAAGTCCCTCA 4233
Oy 298 GlyGlyAlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 316
Db 4234 TCACACAGTACAGCAGCAGCGCAATCCGATTCCAAAGCAGCATCATCATCAAGTCA 4293
Oy 317 AlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValAlaGlnAlaAlaAla 336
Db 4294 TCACACAGTACAGCAGCAGCGCAATCCGATTCCAAAGCAGCATCATCATCAAGTCA 4350
Oy 337 AlaIleAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAla 356
Db 4351 ATATCTAAAGTACAAAGCAATCAGTTCAACACAGCATCGCATCATTAAGTGTGTTCA 4410
Oy 357 IleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyIleThrMet 376
Db 4411 GAAACGATCGATTCAACAAAGCATCAACACAGCAGAGTCA----- 4458
Oy 377 AlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysTrpValThr 396
Db 4459 -----AAATCAGAAAGTACATCCACATCGCTGCTGACTCAACAGTCA 4503
Oy 397 ValGlyValGlyValAlaValAlaAlaProAlaLeuGlyLys----- 410
Db 4504 AGCACTCAGGATCAGCAAGTACATCATGCTCAGTAATCACTCAGCAAGCCAGTGAA 4563
Oy 411 -----GlyIleMetGlnMetGlnLeuSerGlnMetGlnGlnAsnVal 424
Db 4564 TCCGATTCGCTCAACATCATTAAGTAGATTCACAGCTCAGCTCAATGCAAGCAGTGA 4623
Oy 425 AlaGlnPheGlnLys-----GluValGlyLysLeuGlnAlaAlaAlaAspMetIle 441
Db 4624 TCCGATTCAGCAACATCAACATCAATTAATTCACAAAGTACATCAATCAATTA 4683
Oy 442 SerMetPheThrGlnPheTrpGlnGlnAlaSerLysIleAlaSerLysGlnThrGly 461
Db 4684 CGAATGTGACT-----ATTGCAAGTGAATGAGTTCGGAAGTACATCAGAA 4731
Oy 462 SerAsnGluMetThrGlnLysAlaThrLys 471
Db 4732 TCTGTTCAACATCAAGAAAGTACAGTGA 4761

```

```

KW bactericide; vaccine; ds.
XX
OS Fusobacterium necrophorum.
XX
FH Key Location/Qualifiers
FT CDS 1..9726
FT /tag= a
FT /product= "Leukotoxin"
XX
PN WO200180886-A2.
XX
PD 01-NOV-2001.
XX
PE 25-APR-2001; 2001WO-US13240.
XX
PF 25-APR-2000; 2000US-0558257.
XX
PR 24-APR-2001; 2001US-0841786.
XX
PA (UNIV ) UNIV KANSAS STATE RES FOUND.
XX
PI Nagaraja TG, Stewart GC, Narayanan SK, Chengappa MM;
DR WPI: 2002-049245/06.
DR P-PSDB: AAG66005.
XX
PT Fusobacterium necrophorum polypeptide useful as vaccine in immunizing
PT an animal against an infection e.g. foot rot, or liver abscesses caused
PT by the bacterium -
XX
PS Claim 10; Page 90-95; 108pp; English.
XX
CC The invention provides an isolated Fusobacterium necrophorum leukotoxin
CC polypeptide and its truncated versions. The leukotoxin protein can be
CC expressed by standard recombinant methodology. The leukotoxin gene is
CC useful for preparing a vaccine which confers effective immunity against
CC infection caused by F. necrophorum. The vaccine is useful for immunising
CC an animal against liver abscesses caused by F. necrophorum and for
CC preventing foot rot caused by F. necrophorum infection. The present
CC sequence represents a DNA encoding a F. necrophorum full-length
CC leukotoxin polypeptide.
XX
SQ Sequence 9726 BP; 3678 A; 1304 C; 2238 G; 2506 T; 0 other;
XX
Alignment Scores:
Pred. No.: 0.0152 Length: 9726
Score: 167.50 Matches: 114
Percent Similarity: 37.26% Conservative: 79
Best Local Similarity: 22.01% Mismatches: 227
Query Match: 7.07% Indels: 99
DB: 24 Gaps: 17
US-09-889-314-2 (1-496) x AA167640 (1-9726)
Oy 15 GlnLysAsnIleMetSerGlnValLeuThrSerThrProGlnGlyValProGlnLysAsp 34
Db 6641 CAGCAAAACGTTAATGTTCTGTTCGAATTCGAAATGCGGAGTCGAGTGCAGAAAG 6700
Oy 35 LysLeuSerGlyAsnGlnThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGlu 54
Db 6701 GAGCTGGAATTCGACACAGTCGATACCAAGAAAGAAACCAACAGCAGAGCAGAG 6760
Oy 55 MetGluSerAspAlaThrIleLeuAlaGlyLysSerGlyLysAspLysThrSerSerThrThr 74
Db 6761 TG-AAAAATTCATAAATTAATGACT-----CGAAACAGTTCAGATGTAATACCA 6807
Oy 75 LysThrGlnThrAlaProGlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGln 94
Db 6808 CAAATGAGATTAATACGATGCTGCAATCGGTCA----- 6843
Oy 95 LysAlaGlyAlaAsp-----ThrGlyValSerGlyAlaAlaAlaAlaThrAlaSer 111
Db 6844 ---GCCGAGCTGGAATTCCTTGACAGCCGAGATATCTGGA-----GTGGTTCTGTCAAT 6894

```

Leukotoxin; infection; immunisation; liver abscess; foot rot;

```

Db 8699 GGAGCATG---CGAATGAGAGCTGCTATGAGAAATTAATTCATGGAATCAAT 8755
QY 368 ValPheAlaIysGlyThrGlnMetIleAlaIysAsnPhleProIysLeuSerIysValIle 387
Db 8756 ATCAGTATTAAATTCATGATATGAGAAAATAATGATGTTTGTAAAGATATA 8815
QY 388 SerSerLeuThrSerIysTrpValThrValGlyValGlyValValAlaIleProAla 407
Db 8816 TCGGATTTGAGACGAGCAAGAAAGATTAACCGTAGAGCGGTACCTCCGAGCCATT 8875
QY 408 LeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnIleAsnValAlaGlnIle 427
Db 8876 ATCTCAAAAGCAAAAGATGAATGAAT---TCAGAGCTTGAATTTAGAGAGATTTTC 8932
QY 428 GlnLysGluValGlyLysLeuGlnAlaAlaIleAspMetIleSerMetPheThrGlnPhe 447
Db 8933 AATGAAACA----- 8941
QY 448 TrpGlnGlnAlaSerIysIleAlaSerIysGlnThrGlyGlnSerAsnGlnMetThrGln 467
Db 8942 -----AATGAGTACTAGCCCTTCTAAAGCAATTCGAAGAAATCAATGTC 8989
QY 468 LysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTrpAlaIleSerGlyAlaIle 487
Db 8990 AAGGTGAAAAGCAAAACAGAGTGCCTGTAATCTCAAGAGCTTCTGTAGAGCAGTA 9049
QY 488 AlaGlyAla 490
Db 9050 GCAAGCGCA 9058

RESULT 37
AAS54978
ID AAS54978 standard; DNA: 7035 BP.
XX
XX AAS54978:
XX
XX 13-FEB-2002 (first entry)
XX
XX Staphylococcus aureus DNA for cellular proliferation protein #1290.
XX
XX Antisense: ds; prokaryotic cellular proliferation gene;
XX
XX antibiotic; antibacterial; drug design.
XX
XX Staphylococcus aureus.
XX
XX WO200170955-A2.
XX
XX 27-SEP-2001.
XX
XX 21-MAR-2001; 2001WO-US09180.
XX
XX 21-MAR-2000; 2000US-191078P.
XX
XX 23-MAY-2000; 2000US-206848P.
XX
XX 26-MAY-2000; 2000US-207727P.
XX
XX 23-OCT-2000; 2000US-242578P.
XX
XX 27-NOV-2000; 2000US-253625P.
XX
XX 22-DEC-2000; 2000US-257931P.
XX
XX 16-FEB-2001; 2001US-269308P.
XX
XX PA (ELIT-) ELITRA PHARM INC.
XX
XX Haselbeck R, Ohlsen KL, Zyskind JW, Wall D, Trawick JD, Carr GJ;
XX Yamamoto RT, Xu HH;
XX WPI; 2001-611495/70.
XX P-PSDB; AAU37119.
XX
XX New polynucleotides for the identification and development of
XX antibiotics, comprise sequences of antisense nucleic acids -
XX
XX Claim 27; Seq ID No 8615; 51pp; English.
XX
XX The invention relates to antisense inhibitors of genes essential to

```

prokaryotic cellular proliferation, their use in identifying the genes, their use in the discovery of novel antibiotics, the essential genes themselves and the encoded proteins. The prokaryotes used are *Escherichia coli*, *Staphylococcus aureus*, *Salmonella typhi*, *Klebsiella pneumoniae*, *Pseudomonas aeruginosa* and *Enterococcus faecalis*. The invention is also useful for the identification of potential new targets for antibiotic development. The antisense nucleic acids can also be used to identify proteins used in proliferation, to express these proteins, and to obtain antibodies capable of binding to the expressed proteins. The proteins can be used to screen compounds in rational drug discovery programmes. The antisense nucleic acid sequence is also useful to screen for homologous nucleic acids which are required for cell proliferation in a wide variety of organisms. The present sequence encodes an essential prokaryotic cellular proliferation protein.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pcl_sequences.

CC Sequence 7035 BP; 2534 A; 1382 C; 1316 G; 1803 T; 0 other;

Alignment Scores:

	Pred. No.:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
Score:	0.0104	7035	167.50	101	98	234	
Percent Similarity:	39.02%						
Best Local Similarity:	19.80%						
Query Match:	7.07%						
DB:	23						17

US-09-889-314-2 (1-496) x AAS54978 (1-7035)

```

QY 5 SerIleSerSerSerSerGlyProAspAsnGlnIysAsnIleMetSerGlnValLeuThr 24
Db 3334 AGTGATGCAATCAGCTTCACTGATCATGATGATCAATCAATCAAGTATGACACA 3393
QY 25 SerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrIysGlnIle 44
Db 3394 AGTACATCTGCTTGGTAACTGATCTGACATCTGTAAGCAAGTAAATCAGAACGC-- 3450
QY 45 GlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGlyAla 64
Db 3451 -----ACATCAACATCTATG--AGTCATTCACACAGCTTAAGTACA 3489
QY 65 SerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGlyVal 84
Db 3490 TCAGAGTCTGATTCACAGCAGCACTCAACGCGATCTATCAAGTGAAGGATA 3549
QY 85 AlaAlaGlyLysGlnSerSerSerGlnLysAlaGlyAlaAspThrGlyValSerGly 104
Db 3550 ---TCTGTTTCAGAGAGT-----ACGTCATATATCATTA 3579
QY 105 AlaAlaIleThrThrAlaIleSerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlu 124
Db 3580 AGTAATCAATATGATACAGATGATTCAGAGTCAAGAAAGTGCATCTGTTTAAAGGAA 3639
QY 125 GlnAlaSerLysSer---MetGlnSerThrLeuGlnSerLeu-----Gln 138
Db 3640 TCATTGAGTCAAAAGCATCGATCGAATCAACATCAACATCAAGTGTTCACACAGCAT 3699
QY 139 SerLeuSerAlaAlaGlnMetLysGlnValGlnValValValAlaAlaIleLeuSerGly 158
Db 3700 AGTACGCTTTATATAGATGATGATTCGATCTGTGACACTTCACATCATTAAGTAAT 3759
QY 159 LysSerSerGlySerAlaLysLeuGlnLysThrProGlnLeuProIysProGlyVal---Thr 177
Db 3760 TCACACAGTGTAGTGCCTCATTTCAACATCGACAAAGTGTAGCCGATCAACATCAACG 3819
QY 178 ProArgSerGluValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGln----- 194
Db 3820 GTTAAGACGAGAGTGTCTCAACATCATTTGAGCACATCAACAGATTAAGTTGTCTGAC 3879
QY 195 -----ThrLeuGlyGlnAlaThrLysSerAlaLeuSerAsnTrpAla 208

```

Db 4863 CAATGAAATTCGATTCGATGTAGTGAATCCAGAGAAACGATTCGAATGAAATTCGCG 4922
QY 434 ucGlnAlaAlaAspMetIleSerMetPheThrGlnPheTrpGlnGlnAlaSerLys11 454
Db 4923 TCAGGCGAGGTGGAATGCAACGTCTAGTATGACCAAT-----4959
QY 454 eaLaserLysGlnThrGlyLysSerAsnGluMetThrGlnLysAlaThrLysLeuGlyAl 474
Db 4960 ----AGCAGACACCACTGGCAGATGAACATCAACATATCTCAGCATTCGAGGCCAATGC 5015
QY 474 acGln 475
Db 5016 CAGG 5019
RESULT 36
AA167647
ID AA167647 standard; DNA; 11130 BP.
XX
AC AA167647;
XX
DT 27-FEB-2002 (first entry)
XX
DE F. necrophorum leukotoxin gene sequence.
XX
KW Leukotoxin; infection; immunisation; liver abscess; foot rot;
KW bactericide; vaccine; ds.
XX
OS Fusobacterium necrophorum.
XX
PN WO200180886-A2.
XX
PD 01-NOV-2001.
XX
PE 25-APR-2001; 2001WO-US13240.
XX
PR 25-APR-2000; 2000US-0558257.
PR 24-APR-2001; 2001US-0841786.
XX
PA (UNIV) UNIV KANSAS STATE RES FOUND.
XX
PI Nagaraja TG, Stewart GC, Narayanan SK, Chengappa KM;
DR MPI; 2002-049245/06.
XX
PT Fusobacterium necrophorum polypeptide useful as vaccine in immunizing
PT an animal against an infection e.g. foot rot, or liver abscesses caused
PT by the bacterium -
XX
PS Disclosure; Page 102-107; 108pp; English.
XX
CC The invention provides an isolated Fusobacterium necrophorum leukotoxin
CC polypeptide and its truncated versions. The leukotoxin protein can be
CC expressed by standard recombinant methodology. The leukotoxin gene is
CC useful for preparing a vaccine which confers effective immunity against
CC infection caused by F. necrophorum. The vaccine is useful for immunising
CC an animal against liver abscesses caused by F. necrophorum and for
CC preventing foot rot caused by F. necrophorum infection. The present
CC sequence represents the F. necrophorum full-length leukotoxin gene
CC sequence.
XX
SQ Sequence 11130 BP; 4205 A; 1472 C; 2511 G; 2942 T; 0 other;
XX
Alignment Scores:
Pred. No.: 0 015 Length: 11130
Score: 168.50 Matches: 111
Percent Similarity: 36.38% Conservative: 72
Best Local Similarity: 22.07% Mismatches: 215
Query Match: 7.11% Indels: 105
DB: 24 Gaps: 16
US-09-889-314-2 (1-496) x AA167647 (1-11130)
QY 52 AsnThrGluMetGluSerAspAlaThrIleAlaGlyAla-----64

Db 7673 AACGCAAAACGTTATAGTCTCTGCAANTGGAATGCCGAGTCGAGTCGCAAAA 7732
QY 65 ---SerGlyLysAspLysThrSerThrLysThrGluThrAlaProGlnGlnGly 83
Db 7733 GGAGCTCGAATGGAGAGCAGCAGCTGACCAAAAGATGATCAACACGAGCAAGA 7792
QY 84 ValAlaAlaGlyLysGluSerSerGluSerGln-----94
Db 7793 GTGAAAAATTTCTAAATTAATATGACTCGAAACAGTTAGATGTAAATGCGAAAAATGAGATA 7852
QY 95 LysAlaGlyAlaAspThrGlyAlaSerGlyAlaAlaAlaThrThrAla-----110
Db 7853 AATACAGTACTGGAATGCGGTCACCGCAGCTGGAATCTTCGACCGGAGTATCTGGA 7912
QY 111 -----SerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGluGluAla 126
Db 7913 GTGCTTCTGTCATTAATATTCGAAATAGCTA-----GAAACAGATATCGATCATAGT 7966
QY 127 Ser-----LysSerMetGluSerThrLeuGluSerLeuGlnSerLeuSerAlaGln 144
Db 7967 ACTTTACACTCTTCTACTGATGATTAATGTAAGCTTTAATAAATTTGCAATTCCTTG 8026
QY 145 MetLysGluValGluAlaValAlaValAlaAlaLeuSerGly-----158
Db 8027 ACAGCGCGTGAGAGAGCCGACAGTCTTGACAGAGTTACCGAGTGTTCTGTTAACACT 8086
QY 159 ---LysSerSerGlySerAlaLysLeuGlu-----ThrProGluLeuPro 172
Db 8087 ATTAATAGTTCTGTATGATGCTGATGTCACAAATAACTGATTTGACTCCGTACAGAA 8146
QY 173 LysProGlyValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAla 192
Db 8147 AAAGTAATGTAACGCCAAGAGAGAGAAAA-----AATATTAGCAAAACACGACGA 8197
QY 193 IleGlnThrLeuGlyGluAlaThrLysSerAla-----LeuSerAsnThrAlaSer 209
Db 8198 AATGCAAGAAATCGAGAGCAGACCAATCGAGCCAAATGCTTGTAATATTTTGGACAA 8257
QY 210 ThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGluLysGlnAlaIleLysIle 229
Db 8258 -----GCTGTAGAGATGTGAAAAAATTCGAAAGAGAAAGCAAGAGTTTAAAAACT 8311
QY 230 AspLysGluArgGluGluGlyArgGlnGluMetLysAlaAlaGluGlnLysSerLysAspLeu 249
Db 8312 TTAGACGAAGTTTAACAAAGACAGATTAATAAGTAATGATGCTACGAAAAAATCTTA 8371
QY 249 -----249
Db 8372 CAATCAGCAGTATTTCTACAGACATCTCTGTAAGAACGCGATRAGAGACTCTCAG 8431
QY 250 ---GluGlyThrMetAspThrValAsnThrVal---MetIleAlaValSerValAlaIle 267
Db 8432 GGAGAGGAATTAAGCCATTTGTGAAGACTCTGTGATTTGGAAGAAATCTAATAT 8491
QY 268 ThrValIleSerIleValAlaAlaIlePheThrGlyAlaGlyLeuAlaGlyLeuAla 287
Db 8492 ACAACAGAGGACAGAAATATATCACTTCTACTGCTGTTGGAACTGCAGCTCTGCT 8551
QY 288 AlaGlyAlaAlaValAlaGlyAlaAlaAlaAlaGlyAlaAlaGlyAlaAlaAlaAlaThr 307
Db 8552 TCC-----GCATCAGAAACAGTGCAGCTTACA 8578
QY 308 ThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIle 327
Db 8579 AATATTAAAGAAATTCGCGAGTACTCTGTGAATAATCTTTGTGAAGCACCTGAAAAA 8638
QY 328 ThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyLys 347
Db 8639 GTAAATGTTAGTCGAGATTTACAGAGAAATGTTGCTTTAACGCAATATCAAGGCTCTGTA 8698
QY 348 AlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLysGlyLysSerLys 367

OS Drosophila melanogaster.
XX WO200171042-A2.
PN
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001: 2001WO-US09231.
XX
PR 23-MAR-2000: 2000US-191637P.
PR 11-JUL-2000: 2000US-0614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI: 2001-656860/75.
DR P-PDB: ABB59028.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
XX
PS Claim 1; SEQ ID NO 3875; 21bp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
CC sequences (AB101840-AB16175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 10910 BP; 2875 A; 3179 C; 3157 G; 1699 T; 0 other;
SQ

Alignment Scores:
Pred. No.: 0.0147 Length: 10910
Score: 168.50 Matches: 122
Percent Similarity: 34.69% Conservative: 66
Best Local Similarity: 22.51% Mismatches: 210
Query Match: 7.11% Indels: 145
DB: 23 Gaps: 20

US-09-889-314-2 (1-496) x ABL03131 (1-10910)
QY 8 SerSerSerGlyProAspAsnGlnLysAsnLleMetSerGln---ValLeuThrSerThr 26
DB 3607 AGTTCGGCTAATCCAGCCAGGCTCAGGCCGCCAGCGCATTTATTCGATTCGATG 3666
QY 27 ProGlnGlyVal-----ProGlnGlnAspLysLysSerGlyAsnGlu 40
DB 3667 ACCACTGGCGTTATGACGCGATTGACGCGGACGACAAACAAGTCCGCGGTCTTCA 3726
QY 41 ThrLysGlnLleGlnInThrArgGlnGlyLysAsnThrGluMetGlu----- 56
DB 3727 TCCGCCGTGCTGAGCGAGCGGTGCGAAGCTGCCGCCGCGCTGTGCGACAG 3786
QY 57 -----SerAspAlaThrLleAlaGlyValAserGlyLysAspLysThrSerThr 73
DB 3787 AAACGGCAGCTGATGCTGCCAAAGCGGCTGAAGTGCAGCCGCAAAAGCGACGCGCC 3846
QY 74 ThrLysThrGlnThrAlaProGlnGlnGlyValAlaGlyLysGluSerSerGluSer 93
DB 3847 ACTAAGCTGTCGACGCA-----ACGGCTGCGGTGAAAGGTGCGCAAAAGCG 3894
QY 94 GlnLysAlaGlyAlaAspThrGlyValSerGlyLysAlaAlaThrThrAlaSerAsnThr 113
DB 3895 GCTCAGCAGCGCGC-----GCGGCAAGCAACTACCGCCACACCGCA 3939
QY 114 AlaThrLysLleAlaMetGlnThrSerLleGlnLleAlaSerLysSerMetGluSerThr 133

DB 3940 ACAACACACCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCA 3999
QY 134 LeuGlnSerLeuGlnSerLeuSerLysAlaAlaGlnMetLysGlnValGlnAlaVal 153
DB 4000 GCATCATCA---TCGACGGCAGAGTCCAGCAGCATTTGCGACAGCACTACCTGCAACA 4056
QY 154 AlaAlaLeuSerGlyLysSerGlySerGlyLysLysLysGlnThrProGluLeuProLys 173
DB 4057 GATGCC-----GGCAAAAGGCTAGCGCCAGCAT----- 4086
QY 174 ProGlyValThrProArgSerGluValLleGluLleGlyLeuAlaLeuAlaLysAlaLle 193
DB 4087 -----AGATGCA 4095
QY 194 GlnThrLeuGlnGlnAlaThrLysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGln 213
DB 4096 GCAACAGCTGTGAGCAACGCGCAACGGAATCCGACAGCACTACCTGCAACA 4155
QY 214 AlaAspGlnThrAsnLysLeuGlnGlyLeuGlnAlaLleLysLleAspLysGlnArg 233
DB 4156 GCAACAGCACCCAGCATCTCAGCGCGCGAGGC----- 4194
QY 234 GluGlnThrGlnGlnMetLysAlaAlaGlnGlnLysSerLysAspLeuGlnGly----- 251
DB 4195 -----AAGACCAAAATGCCGAGAGAGGACCTGCCACAGCCGAGCAGCA 4242
QY 252 ThrMetAspThrValAlaSerThrValMetLleAlaValSerValAlaLleThrValLleSer 271
DB 4243 ACGGTGCGCATCGCGGTACACCGCAACCGGAGCAAGTCCGCGC----- 4287
QY 272 IleValAlaLleIlePheThrLysGlyAlaGlyLeuAlaGlyLeuAlaGlyAlaLle 291
DB 4288 -----AGTCTGGCAGGCGCAACGCGACAGCGCGCAGC 4323
QY 292 ValGlyAlaAlaAlaGly-----GlyAlaAlaGlyAla 303
DB 4324 GCGACAGCTCGCGCCAGGCGTGTGCAACCGGAACGACGACTGACGCGGAGACG 4383
QY 304 AlaAla-----AlaThrThrValAla 310
DB 4384 GCTGCTAAGGAGCTGACTCGGCCAGATGCCAAGCATCCGTCGCCAAATCGCTCG 4443
QY 311 ThrGlnLleThrValGln-AlaValValGlnAlaValLysGlnAlaValLleThrAlaVal 330
DB 4444 AAGCTATCATATCGGAGGCTATATGCTATGCTGTCGCAACGACACAGCAGCAGCAGC 4503
QY 330 IArgGln---AlaLleThrAlaAlaLleLysAlaAlaValLysSerGlyLleLysAlaPhe 349
DB 4504 AATGCAACGGGCGCC-TCTGCTCCAGTGCAGGCGGCTTACTTGATGCTTTAAGCTTG 4562
QY 349 eLleLysThrLeuValLysAlaLleAlaLysAlaLleSerLysGly----- 364
DB 4563 CTATCAACCGTGTGATGCGCAATGTAAAGCATCCACGGGTGAGATGCTGCTGC 4622
QY 365 -----IleSerLysValPheAlaLysGlu 372
DB 4623 GAACGCTGTGAGACACACCCCGATGTTAGCAGCAACATCATCATCATCATCATCAT 4682
QY 372 YThrGlnMetLleAlaLysAsnPhePro-----LysLeuSerLysValLleSerSerLle 390
DB 4683 TGGCGATTAATCGTCAAGACACACCATCATTCAGGCGCCCAAAATTCATCTTCAC 4742
QY 390 uThrSerLysThrValThrValGlyValGlyVal-----ValVal 403
DB 4743 GCGCGCAATGAGTGTAGTGTGCGGAGTGAATTAATCATGCGCACACCGCAACAC 4802
QY 403 lAlaAlaProAlaLeuGlyLysGlyLleMetGlnMetGlnLeu----- 417
DB 4803 GCGCGGCAATTAATCTTGACGAGAGCTGAAGCGTCCCAAGTTAGAGGTTTGGAGCTGG 4862
QY 418 -SerGluMetGlnGlnAsnValAlaGlnPheGlnLys-----GluValGlyLysLe 434
DB 434

PI Ison CH, Duggan BM, Sapperstein SK;
XX WPI: 2002-519296/55.
DR P-PSDB; AAO21705.
XX
PT Human secreted proteins and polynucleotides for diagnosing, treating or
PT preventing disorders of cell proliferative, cardiovascular,
PT developmental, neurological and autoimmune/inflammatory disorders -
XX
PS Claim 5; Page 218-220; 229pp; English.

XX The invention relates to an isolated human secreted protein (SECP)
CC polypeptide from 63 fully defined protein sequences given in the
CC specification. The polypeptide is useful for the diagnosing/treating of a
CC disease with decreased/overexpression of SECP. Examples of disorders
CC associated with abnormal expression of SECP include a cell proliferative
CC disorder e.g. arteriosclerosis, cancers; autoimmune/inflammatory
CC disorder e.g. allergies, anaemia, asthma; cardiovascular disease e.g.
CC congestive heart failure, ischaemic heart disease; developmental disorder
CC e.g. renal tubular acidosis, hypothyroidism; neurological disorder e.g.
CC Alzheimer's disease, dementia, Parkinson's disease, epilepsy or stroke.
CC The SECP polynucleotide and polypeptide are further useful for analysing
CC the proteome of a tissue or a cell type. The polynucleotide is useful for
CC creating knockin humanised animals (pigs) or transgenic animals (mice or
CC rats) to model human disease, and for somatic or germline gene therapy,
CC and further for generating hybridisation probes useful in mapping the
CC naturally occurring genomic sequence. This polynucleotide sequence
CC represents the DNA of a human secreted protein of the invention.

XX Sequence 6108 BP: 1313 A; 1838 C; 1971 G; 986 T; 0 other;

Alignment Scores:

Pred. No.: 0.00442 Length: 6108
Score: 171.50 Matches: 108
Percent Similarity: 37.36% Conservative: 59
Best Local Similarity: 24.16% Mismatches: 187
Query Match: 7.24% Indels: 93
DB: 24 Gaps: 23

US-09-889-314-2 (1-496) x AAL39666 (1-6108)

QY 10 SerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeuThrSerThrProGlnGly 29
DB 1714 ACAGGCATGAGCACCACCAAACT-----GTCCTAACGCGTAC----- 1752
QY 30 ValProGlnGlnAspLysLeuSerGlyAsnGluThrLysGlnIleGlnIlnThrArgln 49
DB 1753 -----AAGACACACCTTCTACAGTGGGCTGACCATGCTGAACGTGGCCAAAGGGG 1803
QY 50 GlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGlyAlaSer----- 65
DB 1804 GGTGTGCAAACTGGGTGTAAGAAAGCAAAATATCGCAGACAGTACAAAGAACCTTT 1863
QY 66 GlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGlyAla 85
DB 1864 GGCAGTGGGGTGCCTGCTGTGAATGTGCCAAAGGGGTGTCGACAGACAGCTGTAGAC 1923
QY 86 AlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThrGlyValSerGlyAla 105
DB 1924 ACAGCCAAAGACCGTCTGACCGCACCAAGACACATCATCTGTGGGTGAGGGA 1983
QY 106 AlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlnGlu 125
DB 1984 GTG-----AATGTCCGCAAAAGGACGTGTCCAGACAGTGTGGACACC 2025
QY 126 AlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAlaGlnMet 145
DB 2026 ACCAAGACTGTCTTAAGTGTACCAAGACACCGTCTGCAGT----- 2067
QY 146 LysGluValGluAlaValValAlaAlaLeuSerGlyLysSerSerLysAlaLys 165
DB 2068 -----GGGGTGACCGGTGCTGCGAATGTGGCCAAAGGGGCGTCCACAGCGGT 2115

QY 166 LeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnValIleGlnIle 185
DB 2116 GTAGACACT-----ACAAAGTCTGCTCGACTGTGCACATAAGATGCTGTCCACT 2166
QY 186 GlyLeuAla-----LeuAlaLys---AlaIleGlnThrLeuGlyGlnAlaThr 200
DB 2167 GGGCTCACAGGGGCTGTGAACCTTGGCCAAAGGAGACTGTCCAGACCGCATGGACACACC 2226
QY 201 LysSerAlaLeuSerAsnThrLysThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
DB 2227 AAGACTGTGTAACTGCT---ACCAAGATGCTGTGTGACAGTGGGGTGAACCGTCTCGC 2283
QY 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGluArgGluGlnGlnMetLys 240
DB 2284 AATGTGCCAAAGGGGCGCTCCAG-----ACG 2310
QY 241 AlaAlaGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThr---Val 259
DB 2311 GGTGTAGACACGGCCCAAGACCGTGTGACCGCACCAAGGACACAGTCACTGGCGTC 2370
QY 260 MetIleAlaValSerValAla-----IleThrValIleSerIleValAlaAla 275
DB 2371 ATGGGGGCGATGATATGTGCGCCAAAGGACCGTCCAGACCATGTGTGACACCAAGACT 2430
QY 276 IlePheThr-----CysGlyAlaGlyLeuAlaGlyLeuAlaAlaGly 289
DB 2431 GTCTTAAGCTGTACCAAGGACACCGTGTGC---ATGTGGGTGACCGGT---GCTCGAAT 2484
QY 290 AlaAlaValGlyAlaAlaAlaAlaGly----- 298
DB 2485 GTGGCCAAAGGGGCGCTCCAGGGGCGCTGACACTACAAAGTGTCTGTGACGCACT 2544
QY 299 -----GlyAlaAlaGlyAlaAla-----AlaAlaThrThrValAla 310
DB 2545 AAGACACCGCTTCCACTGTGCTGACAGGGGCTGTGAACCTGGCCAAAGGACGTGTCCAG 2604
QY 311 ThrGlnIle---ThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAla 329
DB 2605 ACCGGCGTGGACACGACGACGACTGTCTGACCGGTACCAAGACACCGTGTGACGTGA 2664
QY 330 ValArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPhe 349
DB 2665 GTCACTGTGTGCGTAAATGTGGCC---AAAGGACCGTCCACAGAGTGTGACACA--- 2718
QY 350 IleLysThrLeuValLysAlaIleAlaLysAlaIleSerLysGlyIleSerLysVal--- 368
DB 2719 GCCAAGACGCTGTGAGTGGCGCTTAAGATGCAATGACTGAGTCAAGGGGCGACGTG 2778
QY 369 ---PheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeuSerLysValIle 387
DB 2779 AATGTGCCCAAAAGAACCGTGCAGACCGGCGTGGAC-----GCCCTCAAGGCTGTG 2829
QY 388 SerSerLeuThrSerLysTrpValThrValGlyValGlyValAlaValAlaPheAla 407
DB 2830 CTTATGGGTACCAAGACACTGTCTTCACTGG-----GTTAACCGGTCCATGAC 2880
QY 408 LeuGlyLysGlyIleMetGln 414
DB 2881 ATGCCAAAGGGGCGTCCAG 2901
RESULT 35
ABL03131
ID ABL03131 standard; cDNA; 10910 BP.
XX
AC ABL03131;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 3875.
DE
XX Drosophila: developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX

DR WPI: 1998-272225/24.

XX Computer-readable medium with recorded Streptococcus pneumoniae

PT polynucleotide sequences - useful in diagnostic kits and assays, and

PT pharmaceutical compositions and vaccines for Streptococcus

XX pneumoniae

PS Claim 1: Page 1384; 1409pp; English.

XX

XX The present invention describes a computer readable medium which has

CC the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)

CC recorded on it, or a representative fragment or a sequence at least 95%

CC identical to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in

CC SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from

CC Streptococcus pneumoniae. The present invention also describes an

CC isolated nucleic acid molecule encoding a homologue of any of the

CC fragments of the S. pneumoniae genome (SEQ ID NO:1 to 391) where the

CC nucleic acid molecule is produced by a process comprising: (a) screening

CC a genomic DNA library using as a probe a target sequence defined by any

CC of the sequences in SEQ ID NO:1 to 391, identifying members of the

CC library which contain sequences that hybridize to the target sequence and

CC isolating the nucleic acid molecules from the members; or (b) isolating

CC mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid

CC molecules whose nucleotide sequence is homologous to amplification

CC primers derived from the fragment of the S. pneumoniae genome to prime

CC the amplification and isolating the amplified sequences. The computer

CC readable medium can be used in a computer-based system for identifying

CC fragments of the S. pneumoniae genome of commercial importance, or

CC expression modulating fragments of the S. pneumoniae genome. Products

CC from the present invention can be used in the diagnosis kits and assays, and

CC pharmaceutical compositions and vaccines for S. pneumoniae.

XX

SQ Sequence 869 BP: 211 A; 277 C; 211 G; 169 T; 1 other;

Alignment Scores:

Pred. No.:	Length:	869
Score: 172.50	Matches: 74	
Percent Similarity: 41.14%	Conservative: 49	
Best Local Similarity: 24.75%	Mismatches: 138	
Query Match: 7.28%	Indels: 39	
DB: 19	Gaps: 6	

US-09-889-314-2 (1-496) x AAV52502 (1-869)

OY 59 AlAThrIleAlaGlyAlaSerGlyLysAspLysThrSerThrThrLysThrGluThr 78

DB 6 AGTACTAGTGCATCAGCTTCAGCATCAAGAGTGCATCGCTTCGCTCAACAGTCA 65

OY 79 AlAProGlnGlnGlyAlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAla 98

DB 66 TCAGAGTCACAGTACAGTACAGTGCCTTCAGCATCAACAGTGCCTTCAGCA 125

OY 99 AspThrGlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAla---ThrLysIle 117

DB 126 AGCACCGATCGCTCGCTTCAGCAAGTACTAGCGCTTCAGCTCAACAGTCA 185

OY 118 AlAlaMetIleThrSerIle-----GluGlnAlaSerLysSerMetGluSerThrLeu 134

DB 186 TCAGCTCAGCAGCATCAAGTGCCTTCAGCATCAACAGTGCCTTCAGCA 245

OY 135 GluSerLeuGlnSerLeuSerAlaAlaGlnMetLysGluValGlnAlaValAlaValAla 154

DB 246 AGTACTAGTGCCTTCAGCTTCAGCTCAACAGTGCATCGCTTCAGCAAGTGC 305

OY 155 AlAlaSerGlyLysSerGlySerAlaLysLeuGluThrProGluLeuProLysPro 174

DB 306 TCTGAATCGGATCAACAGTGCCTTCAGCATCAACAGTGCCTTCAGCA 350

OY 175 GlyAlaThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGln 194

DB 351 -----GCCCTAGGCTTCAGCTTCACA 371

OY 195 ThrLeuGlyAlaThrLysSerAlaLeuSerAsnTyraLaseThrGlnAlaGlnAla 214

DB 372 AGTGCATCGGCTTCAGCATCAACAGTGCCTTCAGCAAGTGCCTTCAGCA 431

OY 215 AspGlnThrAsnLysLeuGlyLeuGlnLysGlnAlaIleLysIleAspLysGlu----- 232

DB 432 TCAGCTCAACAGTGCATCAGCTTCAGCTCAACAGTGCCTTCAGCAAGTGC 491

OY 233 ----ArgGluGluThrGlnGlnGlnMetLysAlaAlaGlnGlnLysSerLysAspLeuGlu 251

DB 492 AGTGCATCGGCTTCAGCAAGTGCCTTCAGCAAGTGCCTTCAGCAAGTGCCTTCAG 550

OY 251 yThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSe 271

DB 551 CTCAGCATGCACAGGCTTCAGCTTCAGCAAGTGCCTTCAGCAAGTGCCTTCAG 596

OY 271 rIleValAlaAlaIleThrThrCysLysAlaGlyLeuAlaGlyLeuAlaAlaIleAla 291

DB 597 -----GCCCTCAGCTTCAGCAAGTGCCTTCAGCAAGTGCCTTCAGCAAGTGCCTTCAG 649

OY 291 aValGlyAlaAlaAlaAlaGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 310

DB 650 AGCAAGTACCAAGTGCCTTCAGCTTCAGCAAGTGCCTTCAGCAAGTGCCTTCAG 709

OY 311 -----ThrGlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaVal 326

DB 710 TGAATCGGATCAACAGTGCCTTCAGCTTCAGCAAGTGCCTTCAGCAAGTGCCTTCAG 769

OY 326 lIleThrAlaValAlaGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSer 344

DB 770 AAGAGTGCATCGCTTCAGCAAGTGCCTTCAGCAAGTGCCTTCAGCAAGTGCCTTCAG 824

RESULT 34

AAI39666

ID AAI39666 standard; DNA; 6108 BP.

AC AAI39666;

DT 05-SEP-2002 (first entry)

XX

DE Human secreted protein DNA SEQ ID No 110.

XX

KW Antiarteriosclerotic; cytoskeletal; HIV; anti-allergic; anti-naemic;

KW anti-asthmatic; cardiatic; vasotropic; neuroprotective; nootropic; SECP;

KW anticonvulsant; antiparkinsonian; cerebroprotective; anti-inflammatory;

KW immunosuppressive; human secreted protein; cell proliferative disorder;

KW arteriosclerosis; cancer; autoimmune; inflammatory disorder; AIDS;

KW allergy; anaemia; asthma; cardiovascular disease; developmental disorder;

KW ischaemic heart disease; congestive heart failure; neurological disorder;

KW renal tubular acidosis; hypothyroidism; Alzheimer's disease; dementia;

KW Parkinson's disease; epilepsy; stroke; knock in humanised animal;

KW transgenic animal; gene therapy; gene; ds.

XX

OS Homo sapiens.

PN WO200238602-A2.

PD 16-MAY-2002.

XX

PF 08-NOV-2001; 2001MO-US47420.

XX

XX 08-NOV-2000; 2000US-247505P.

PR 09-NOV-2000; 2000US-248642P.

PR 16-NOV-2000; 2000US-249824P.

PR 21-NOV-2000; 2000US-252824P.

PR 08-DEC-2000; 2000US-254305P.

PR 18-DEC-2000; 2000US-256448P.

XX

PA (INCY-) INCYTE GENOMICS INC.

XX

PI Yue H, Yao MG, Gandhi AR, Baughn MR, Swarnakar A, Walla NK;

PI Sanjanwala M, Thornton M, Elliott VS, Lu Y, Gietzen KJ, Burford N;

PI Ding L, Hafalia A, Tang YT, Bandman O, Warren BA, Honchell CP;

PI Lu DM, Thangavelu K, Lee S, Xu Y, Yang J, Dai PG, Tran B;

CC vascular growth. Polypeptides involved with tissue regeneration and
 CC repair (or nucleic acids encoding them) may be used to promote wound
 CC healing (e.g., of burns, incisions and ulcers), while those with
 CC immunomodulatory activities may be used in the treatment of viral,
 CC bacterial and fungal infections in addition to immune disorders.
 CC Polypeptides with growth factor activity may be used in cell cultures to
 CC promote cell growth. For example, such polypeptides may be used to
 CC manipulate stem cells in culture to give rise to neuroepithelial cells
 CC that can be used to augment or replace cells damaged by illness,
 CC autoimmune disease or accidental damage. The polypeptides and nucleotides
 CC may also be used in the diagnosis of the above conditions, and in drug
 CC screening techniques. The present sequence represents a cDNA encoding a
 CC novel human polypeptide of the invention.

CC
 CC Sequence 3750 BP: 873 A; 1057 C; 1222 G; 598 T; 0 other;

Alignment Scores:

Pred. No.:	0.00148	Length:	3750
Score:	174.50	Matches:	107
Percent Similarity:	36.77%	Conservative:	57
Best Local Similarity:	23.99%	Mismatches:	191
Query Match:	7.37%	Indels:	91
DB:	22	Gaps:	21

US-09-889-314-2 (1-496) x ABA09498 (1-3750)

OY 10 SerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeuThrSerThrProGlnGly 29
 DB 1900 ACAGGCAAGGACCAACCACT-----GTCCTAACCGGTACC----- 1938
 OY 30 ValProGlnAspLysLeuSerGlyAsnGlnThrLysGlnIleGlnIleThrArgIn 49
 DB 1939 -----AAGACACCATCTACAGTGGGGTCACCAAGCCGTGACACTGCCCAAGGG 1989
 OY 50 GlyLysAsnThrLysMetLysSerAspAlaThrIleAlaGlyAlaSer----- 65
 DB 1990 GCTGTGCAAACTGGGCTGATAAAGCCAAATATCGCACAGTACCAAGAACCACTTT 2049
 OY 66 GlyLysAspLysThrSerSerThrThrLysThrGlnIleAlaProGlnIleValAla 85
 DB 2050 GGCAGTGGGGTACCAAGTGTGTGTAATGTGGCCAAAGGGGCTGCCAGACAGGTGAGAC 2109
 OY 86 AlaGlyLysGlnSerSerLysSerGlnLysAlaGlyAlaAspThrGlyAlaSerGlyAla 105
 DB 2110 ACGGCAAGACCGCTGCTGACCGCAAGACAGACAGTACACTGCGTCAATGGGGCA 2169
 OY 106 AlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlnIu 125
 DB 2170 GTG-----AATGTGCGCAAGGACACTGTCACAGCAGTGTGACACC 2211
 OY 126 AlaSerLysSerMetLysThrThrLysGlnSerLysLeuGlnSerLysAlaGlnMet 145
 DB 2212 ACCAAGACTGCTTAACTGCTGACCAAGACACCGTCTGCAGT----- 2253
 OY 146 LysGlnValGlnAlaValAlaAlaAlaLysSerGlyLysSerSerGlySerAlaLys 165
 DB 2254 -----GGGGTGAACCGCTGCTGCAATGTGGCCAAAGGGCCATCCAGGGGGC 2301
 OY 166 LeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnValIleGlnIle 185
 DB 2302 CTGGACACT-----ACAAAGTCTGTCTCGCTGCGACATCAAGATGCTGTCCACT 2352
 OY 186 GlyLeuAla-----LeuAlaLys-----AlaIleGlnThrLeuGlnIleGlnAlaThr 200
 DB 2353 GGGCTCACAGGGCGCTGTGAAGTGGCCAAAGGACACTGTCCAGACCGGACGACACACC 2412
 OY 201 LysSerAlaLeuSerAsnThrLysThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
 DB 2413 AAGACTGTGTAACTGCT---ACCAAGATGCTGTGTGACAGTGGGGTACCGGTGCTGG 2469
 OY 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnIuThrGlnIleLys 240
 DB 2470 AATGTGCGCAAGGGGCGCTCCAGATG----- 2496

OY 241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrVal----- 256
 DB 2497 GGTGTAGACAGCGCCAAAGACCTGCTGACCGGTACCAAGACACTGTGTCCAGTGGGTC 2556
 OY 257 -----AsnThrValMetIleAlaValSerValAlaIleThrValIleSerIle 272
 DB 2557 ACCGGTGTGCGCAAGCTGGGCCAAGGGTCTGTGCAAACTGGCGTAAAGCCCAAAAT 2616
 OY 273 ValAla-----AlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAla 290
 DB 2617 ATCGCAACAGGTACCAAGAACACCTTGCACTGGGTGGTACCGGT---GTCGCCAAAGTG 2673
 OY 291 AlaValGlyAlaAlaAlaAlaGly----- 298
 DB 2674 GCCAAAGGGGCGCTCCAGGGGCGCTGCACACTACAAAGTCTGTCTGACGTGCACTAA 2733
 OY 299 -----GlyAlaAlaGlyAlaAla-----AlaAlaThrThrValAlaThr 311
 DB 2734 GATGCCGTGTCCACTGTGGCTCACAGGGCTGTGAACTTGCCAAAGGACCTGTCCAGAC 2793
 OY 312 GlnIle---ThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaVal 330
 DB 2794 GCGGTGACACCGACAGACTGTCTGACCCGCTACCAAGGACCGCTGTGACGTGAGTC 2853
 OY 331 ArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyLysAlaPheIle 350
 DB 2854 ACTGTGCTCCGTAAATGTGCT---AAAGGACCTGTCCAGACAGGTGTGACACA---GCC 2907
 OY 351 LysThrLeuValLysAlaIleAlaLysAlaIleSerLysGlyLysSerLysVal----- 368
 DB 2908 AAGACGTGTCTAGTGGCGCTTAAGATGACAGTCACTGACAGTCAAGGGGCGACGTGAT 2967
 OY 369 PheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeuSerLysValIleSer 388
 DB 2968 GTGGCCAAAGAACCGTGCAGACCGCGGTGAC-----GCCTCCAAAGCTGTGCTT 3018
 OY 389 SerLeuThrSerLysThrValThrValGlyValAlaValAlaIleAlaProAlaLeu 408
 DB 3019 ATGGGTACCAAGACACTGTCTTCACTGGG-----GTTACCGGTGCCATGACATG 3069
 OY 409 GlyLysGlyIleMetGln 414
 DB 3070 GCCAAAGGGGCGGTCCAG 3087
 RESULT 33
 AAV52502
 ID AAV52502 standard; DNA; 869 BP.
 AC AAV52502;
 XX
 DT 23-OCT-1998 (first entry)
 XX
 DE Streptococcus pneumoniae genome fragment SEQ ID NO:369.
 XX
 KW Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
 XX computer readable medium; vaccine; pharmaceutical composition; ds.
 OS Streptococcus pneumoniae.
 XX
 PN W09818931-A2.
 XX
 PD 07-MAY-1998.
 XX
 PF 30-OCT-1997; 97WO-US19588.
 XX
 PR 31-OCT-1996; 96US-0029960.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;
 PI Kunsch CA, Rosen CA;
 XX

Db 3008 GTTCAGACTTCTCTATGCCCTTAAGAAATCC-----AGTCTACATGCTGCTTTAACT 3061
 QY 202 SerAlaIeuserAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu--- 220
 Db 3062 TCCGCTTTTCCCGCATGGGAACAAAGAGCAATTTGGACACACGCTTATATAATTGATG 3121
 QY 221 -----GlyLeuGluLysGlnAlaIleLysIleAspLysGluuAGLuglu 235
 Db 3122 AAGAAAAACAATTGACACATTTGAAGAGCCCTGATTGTAAGAGAAAAAGAGAAAG 3181
 QY 236 TyrGlnGlu-----MetLysAlaAlaGlnGlnLysSerLysAspLeuGlnGly 251
 Db 3182 TTGGAGAGAGTTCGTTCTCGGCTTATCGGCTGCTTAAGAGAAAAAGACCTTCCTGACCCA 3241
 QY 252 ThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThr----- 268
 Db 3242 TCTCTGCCATCATCAAAAGAAACGACAGATCTCCCAACGCGATCCGACACAGAGAAAGCTCG 3301
 QY 269 -----ValIleSerIleValAla 274
 Db 3302 CTTCATATCTGTCACACGACGACGTCAAATCAGAAACCTTTACTATCCTGTTACCACA 3361
 QY 275 AlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaIleAlaValAlaGlyAla 294
 Db 3362 GTGCTGGAATAATCCAGAGCGCTCTCTCCACACGCGATCCGACACAGAGAAAGCTCG 3421
 QY 295 AlaAlaIleAlaGlyAlaAlaGlyAlaAlaAlaAlaAlaThrThrValAlaIleThr 314
 Db 3422 TCGCGCAGCGCGGAGGCGCGCGCGCGCGCTTG----- 3457
 QY 315 ValGlnAlaValAlaGlnAlaValLysGlnAlaValIlePheThrAlaValArgGlnAlaIle 334
 Db 3458 -----AGCGCCCTGCGAATATGCGAATGGAGGCTGTGTCATC---AAGACAGTCTC 3508
 QY 335 ThrAlaAlaIleLysAlaAlaValLysSerGlyLysAlaPheIleLys----- 351
 Db 3509 AATGCCCTGTGGCAGACAGCAATAGTGCCACGCCCTCAGACGCTCTCAAAACAACACACCA 3568
 QY 352 ThrLeuValLysAlaIleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLys 371
 Db 3569 AAGATGATGAGTTCCTGCCCCCGCTCMAACGAAAGCTTATGAAACATGCTTAGCGAG 3628
 QY 372 GlyThrGlnMetIle-----AlaLysAsnPhePro-----Lys 382
 Db 3629 ATGCAACAGCGCATGAGATGATCAAGTTTCAACAACTCACACACAGAGCTCAAG 3688
 QY 383 LeuSerLysValIleSerSerLeuThrSer-LysTyrValThr--ValGlyValGlyAl 401
 Db 3689 GTCAACAAGTGGCTTTCAGACATGACAGTCCACTGGGTGACAGCTGAGCATTTGATTAC 3748
 QY 402 ValValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGln 421
 Db 3749 GTAGAGAGAGCGGGGAGAGTGTTCGGGAAGTGGCAATTAATCGCGCTCAACCGCCAG 3808
 QY 422 GlnAsnValAlaGlnPheGlnLysGlnValLysLysGlnAlaAlaAlaAspMetIle 441
 Db 3809 CAGGAAACTGCACAACTCTCAAGTGCAGCCCAATTTGCAAAAGCAACACACAAACACAG 3868
 QY 442 SerMetPheThrGlnPheThrGlnAlaSer----- 452
 Db 3869 CAGCAGACGACCAACAACACACACACGATGCCCGGGCCACAGAAATTTGACAGCA 3928
 QY 453 -----LysIleAlaSerLysGlnThrGlyLys-----SerAsnGlnMet 465
 Db 3929 GAGGAGCCCGTACACAGTATCAATAAAGACAGCAACCGTTGGAGGCAATTAACCG 3988
 QY 466 ThrGlnLys-----AlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyrAla 481
 Db 3989 CCGCAGCT-AAAGAGACTTATGCTTACGAGAAACCCCAATTCAGAAAGTGGATCC 4047
 QY 482 AlaIleSerGlyAlaIleAlaGlyAlaHis 491
 Db 4048 AAAGTGGCTGGAGCAGATGCAAAATCCCAT 4077

RESULT 32
 ABA09498 standard; cDNA, 3750 BP.
 ID ABA09498
 AC ABA09498;
 XX 11-JAN-2002 (first entry)
 DE Human S3-12 homologue-encoding cDNA, SEQ ID NO:1274.
 XX Human; cytokine; cell proliferation; cell differentiation; growth factor;
 KW haematopoiesis regulation; tissue growth; immunomodulator; activin;
 KW inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;
 KW proliferation; metastasis; cancer; tumour; haematopoietic disorder;
 KW myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;
 KW chronic inflammatory condition; proliferative retinopathy;
 KW atherosclerosis; coronary heart disease; arterial ischaemia;
 KW bone disorder; osteoporosis; vascular growth disorder;
 KW tissue regeneration; wound healing; infection; immune disorder;
 KW cell culture; drug screening; gene therapy; antiinflammatory;
 KW antisthmatic; antiarthritic; haemostatic; antiarteriosclerotic;
 KW cytostatic; osteopathic; vasotropic; cardiac; virucide; antibacterial;
 KW antifungal; vulnereary; antitumor; ss.
 XX Homo sapiens.
 OS
 XX WO200157188-A2.
 XX
 PD 09-AUG-2001.
 XX
 PF 05-FEB-2001; 2001WO-US03800.
 XX
 PR 03-FEB-2000; 2000US-0496914.
 PR 27-APR-2000; 2000US-0560875.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Tang YT, Liu C, Drmanac RT;
 XX
 DR WPI: 2001-457740/49.
 DR P-PSDB; ABB12254.
 XX
 PT Human proteins and DNA encoding sequences useful for preventing,
 PT treating or ameliorating a medical condition in a mammalian subject
 PT e.g. arthritis and cancer -
 XX
 PS Claim 1; Page 968-969; 1963pp; English.
 CC Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and
 CC sequences ABA08225-ABA09574 represent nucleic acids encoding them. The
 CC invention also relates to vectors and recombinant host cells comprising a
 CC nucleotide of the invention, methods of producing the novel polypeptides,
 CC antibodies against the polypeptides, methods of detecting the nucleotides
 CC or polypeptides in a sample, and methods of identifying compounds which
 CC bind to polypeptides of the invention. Although novel, many of the
 CC polypeptides of the invention have homology to known proteins, thereby
 CC giving an insight into their probable biological activities, and hence
 CC potential therapeutic applications. The polypeptides of the invention may
 CC have various activities, including cytokine, cell proliferation or cell
 CC differentiation activities; stem cell growth factor activity;
 CC haematopoiesis regulatory activity; tissue growth activity;
 CC immunomodulatory activity; activin- or inhibin-related activities;
 CC chemotactic or chemokinetic activities; haemostatic, thrombotic or
 CC thrombolytic activities; receptor or ligand activities; or may be
 CC involved in oncogenesis, cancer cell proliferation or metastasis.
 CC Depending on their biological activities, polypeptides and nucleotides of
 CC the invention are useful for preventing, treating or ameliorating medical
 CC conditions, e.g., by protein or gene therapy. Such conditions include
 CC cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell
 CC disorders), chronic inflammatory conditions (e.g., asthma or arthritis),
 CC proliferative retinopathy, atherosclerosis, coronary heart disease,
 CC arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal

```

Db 2350 -----TCAGCCAGCAGCAAGTCTCCACAAGTGAAGTACTAGCCCTCA 2394
OY 256 ValaInThValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAla 275
Db 2395 ATGAGTGCCTCGACGTCACGACGACGACGACGACGACGACGACGACGACGACGACG 2454
OY 276 IlePheThrCysGlyValAglCysLeuAglCysLeuAlaAglValAlaValAglAlaAla 295
Db 2455 ATGAGTGCCTCGACGTCACGACGACGACGACGACGACGACGACGACGACGACGACG 2514
OY 296 ---AlaAlaIleGlyValAlaAglValAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 314
Db 2515 ACCAGTGCCTCAACGACGTCGATGTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2574
OY 315 ValGlnAlaValAlaValGlnAlaValIleSerGlnAlaValIleThrAlaValArgGlnAlaIle 334
Db 2575 ATGTCAGCGACGACCAAGTCTTCAACCAAGTGCCTCGATGTCAGCAGCAGCAGCAGCAG 2634
OY 335 Thr---AlaAlaIleLeuAlaValAlaValIleSerGlyIleLeuAlaPheIleLeu 353
Db 2635 ACAAGTGCCTCGACGTCACGACGACGACGACGACGACGACGACGACGACGACGACG 2694
OY 354 ValIleAlaIleAlaValAlaIleSerIleGlyIleSerIleValAlaPheAlaValIle 373
Db 2695 ACCAGCGTCTCAACGACGTCATGACATGACGACGACGACGACGACGACGACGACGAC 2754
OY 374 GlnMetIleAlaValAlaValPheProIleSerIleValIleSerIleThrIleSerIle 393
Db 2755 AGCTGACGACCTTCAATTCATCAAA---GAGAAGTGTATTCCTTACCTCTCT--- 2808
OY 394 TrpValIleThrValGlyValGlyValValValAlaAlaProAlaLeuGlyIleGlyIleMet 413
Db 2809 -----ACGGGAGACCAAGATTATTCGTGTAAGTCTGCTAGCTTACGTTAATG 2862

RESULT 31
ABL06539
ID ABL06539 standard; cDNA; 8199 BP.
XX
AC ABL06539;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster expressed polynucleotide SPQ ID NO 14099.
XX
KM Drosophila: developmental biology; cell signalling; insecticide;
XX
KM pharmaceutical; gene; ss.
XX
OS Drosophila melanogaster.
XX
PN WO2001/1042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR MPI: 2001-656860/75.
XX
DR P-PSDB; ABB62436.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX
PT genes from Drosophila and for elucidating cell signalling and cell-cell
XX
PT interactions -
XX
PS Claim 1; SEQ ID NO 14099; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent

```

```

CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL01840-ABL16175), expressed DNA
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pcl_sequences.
XX
SQ Sequence 8199 BP; 2477 A; 1928 C; 1979 G; 1815 T; 0 other;
XX
Alignment Scores:
Pred. No.: 0.00206 Length: 8199
Score: 178.00 Matches: 129
Percent Similarity: 37.58% Conservative: 101
Best Local Similarity: 21.08% Mismatches: 216
Query Match: 7.51% Indels: 167
DB: Gaps: 24

US-09-889-314-2 (1-496) x ABL06539 (1-8199)
OY 15 GlnIleAsnIleMetSerGlnValIleThr-----SerThrProGlnGly 29
Db 2342 CAGACCTCCACATGACGACGACGACGACGACGACGACGACGACGACGACGACG 2401
OY 30 ValProGln-----GlnAspIleSerGlyIleGlnIleThr 41
Db 2402 GTTCCCGGACGACCCCAATCAAGCCCTTCTTCCACAGTTCACGTTGATGATGCT 2461
OY 42 LysGlnIleGlnGlnThrArgGlnGlyIleAsnThrGlnMetGlnSerAspAlaThrIle 61
Db 2462 AAAAGCATC-----AATGCTGCTGTGACGACGACGACGACGACGACGACGACGAC 2500
OY 62 AlaGlyAlaSerGlyIleAspIleThrSerSerIleThrIleThrGlnIleAlaProGln 81
Db 2501 GCTGCAGCAGGTGGA-----AATGCCACCTCCATCTGCTGCTGCCGGA 2548
OY 82 GlnGlyValAlaIleGlyIleGlnSerGlnSerGlnIleGlyIleAlaAspThrGly 101
Db 2549 ATGTCACGTGACGACCAACTGCTCTTCG-----GCCGGT-----GGA 2587
OY 102 ValSerGlyAlaAlaIleAlaThrThrIleSerAsnThrAlaThrIle----- 116
Db 2588 GTGAGCGATTGAGGCTTCGAGTCCGCGGCTGTGAGTGGAAAGACGCAAGCGACAC 2647
OY 117 IleAlaMetGlnThrSerIleGlnIleAlaSer----- 127
Db 2648 ATGCCCATTTGAGTGGAGGAAAGGCTTAAGCTCCACGCTCTGCTTAACAGTTCTACA 2707
OY 128 -----LysSerMetGlnSerThrIleGlnSerIleGlnSerIleSerIleAlaGln 144
Db 2708 ATGGCGCCGAAAGACTGGAGAGTGAATTGCAACATGAGAGCCCTTGCTGCTCT 2767
OY 145 -----MetIleGlnValGlnAlaVal 151
Db 2768 GGTGCTGTCGAGTGGCAATTTCTGCTTGAGTGGCATCAGCCGCTCAGCTCAT 2827
OY 152 ValValAlaAlaLeuSerGlyIleSerSerGlySer-----AlaIle 165
Db 2828 CAACATGCTTCACTCTCGGACCAATCATCTGACAGATTGACAGACGACGTCGCCGA 2887
OY 166 LeuGluThrPro-----GluLeuProIle 173
Db 2888 ATGAAGACCTCTCCCTCCATGAGTGCACCAATGACCTTTCATCAAGCCTGCCGA 2947
OY 174 ProGlyValThrProArg-----SerGlu 181
Db 2948 ATGATATATGACTGAGATGCTTAAGTCGCTTCCAGTTCGGAGCCATGATTTGAGTGAG 3007
OY 182 ValIleGlnIleGlyLeuAlaLeuAlaIleGlnIleThrIleGlnIleGlnIleThrIle 201

```


PD 14-FEB-2002.
 XX 08-AUG-2001; 2001WO-NZ00160.
 XX 08-AUG-2000; 2000US-0634238.
 PR 28-NOV-2000; 2000US-0724623.
 XX (GENE-) GENESIS RES & DEV CORP LTD.
 PA (VIAL-) VIALACTIA BIOSCIENCE NZ LTD.
 XX Glenn M, Havukkala J, Bloksberg LN, Lubbers MW, Dekker J;
 PI Christenson AC, Holland R, O'toole PW, Reid JR, Coolbear T;
 DR WPI: 2002-241760/29.
 P-PSDB: AAE20110.
 XX New polynucleotides and polypeptides from *Lactobacillus rhamnosus*,
 PT useful in e.g. improving the flavor, aroma, texture and health-related
 PT benefits of milk-derived products, or in increasing properties of
 PT microbes -
 XX
 PS Claim 2: Fig 63; 257pp; English.
 XX
 CC The present invention relates to a new isolated polynucleotide comprising
 CC a sequence present in *Lactobacillus rhamnosus* strain HN001 and encoding a
 CC polypeptide capable of modifying the flavour, aroma, texture, nutritional
 CC and health benefits of milk-derived products, and/or survivability of
 CC microbes in dairy manufacturing processes. The polynucleotides are useful
 CC for improving the properties of microbes used in the manufacture of milk-
 CC derived products such as cheeses, yogurt, fermented milk products, sour
 CC milks and buttermilk; in modifying the flavour, aroma, texture and health
 CC -related benefits of milk-derived products and in increasing the survival
 CC of microbes during industrial fermentation processes. The bacteria may be
 CC used to increase resistance to enteric pathogens and anti-infection
 CC activity, including treatment of rotavirus infection and infantile
 CC diarrhoea; aid in lactose digestion; as anti-cancer and anti-mutagenesis;
 CC liver cancer reduction; reduction of small bowel bacterial overgrowth;
 CC immune system modulation and treatment of autoimmune disorders and
 CC allergies; treatment of allergic responses to foods; reduction of blood
 CC lipids and prevention of heart disease; antihypertensive effect;
 CC prevention and treatment of urogenital infections, *Helicobacter pylori*,
 CC or hepatic encephalopathy; treatment of inflammatory bowel disorder and
 CC irritable bowel syndrome; modulation of endocarditis; and for improved
 CC protein and carbohydrate utilization and conversion. The transgenic
 CC microbial population can be administered to a mammal as an anti-
 CC carcinogenic agent. The present sequence is *Lactobacillus rhamnosus*
 CC outer membrane protein rompa gene.
 XX
 SQ Sequence 4645 BP; 1279 A; 1062 C; 1261 G; 1043 T; 0 other;
 XX
 Alignment Scores:
 Pred. No.: 0.000745 Length: 4645
 Score: 180.00 Matches: 128
 Percent Similarity: 35.588 Conservative: 73
 Best Local Similarity: 22.65% Mismatches: 266
 Query Match: 7.60% Indels: 98
 DB: 24 Gaps: 20
 US-09-889-314-2 (1-496) x AAD31881 (1-4645)
 QY 1 AspThrAsnMetSerIleSerSerSerGlyProAspAsnGlnIlyAsnIleMetSer 20
 DB 2054 GACGGTAATTGGATTACAGACTGCCGACTGGGCTTTCGTACTCCCAATGCTTCATT 2113
 QY 21 GlnValLeuThrSer----- 25
 DB 2114 GAAGTTATTTCAGTCTCGATTTCATTTGAAAAACGACGAGCTACTGATCAACGATGCC 2173
 QY 26 ---ThrProGlnGlyValProGlnGln-----AspIlyLeuSerGlyAsnIlyThrIly 42
 DB 2174 GAACGCCCAAGACGATCCAGCTGCGAGCTTATTCACGCCCAATTAAGTGTGCT 2233
 QY 43 GlnIleGlnGlnThrArgGlnGlyIlyAsnThrGlnMetGlnIlySerAspAla----- 59

DB 2234 GATGTCACAGCTTCACAGCAAAAGCTACAAAGTCTGCTGTAGTATGATCGCGAGATTAT 2293
 QY 60 -----ThrIleAlaGly-----AlaSerGlyIlyAspIlyThrSerSer 72
 DB 2294 GCAGGTGAAGCCCAATCGATTGTCGACATCATGCTGATATATATGAAATCAAGTCTCTC 2353
 QY 73 ThrThrIlyThrGlnThrAlaProGlnGlnIlyValAlaAlaGlyIlyGlnIlySerGlu 92
 DB 2354 GCCAGTATGCTGAGACAGCAATTCGAATTCCTTGGCAGCTAGCAAGTCTCTCGGCT 2413
 QY 93 SerGlnIlyAlaGlyAlaAspThrGly---ValSerGlyAlaAlaAlaThrThrAlaSer 111
 DB 2414 ACTTCACAGCGCGAGCGCCGACCAATCTGCGCAAGTGAAGCGCGGTATGACAGCTCA 2473
 QY 112 AsnThrAlaThrIlyIleAla---MetGlnThrSerIleGlnIlyAlaSerIlySerMet 130
 DB 2474 TCTGCAGCTGCTCCGATGATGACCTGATGATGACCAAACTCTGACCCCTCTTAT 2533
 QY 131 GluSerThrLeuGlnIlySerLeuGlnIlySerAlaAlaGlnMetIlyGlnIlyAla 150
 DB 2534 GATTCCTACGCTTCGAGCGCCAGTCCGCTTCTGCTGATGATGATGATGCGGATATGCC 2593
 QY 151 ValVal---ValAlaAlaLeuSerGlyIlySerSerGlySerAlaIlyLeuGlnIlyThrPro 169
 DB 2594 ACTGCATCATTTGACAGCAAGTTCGCTGCGCTGCATGACGAGCGATTATGACAGCG 2653
 QY 170 GluLeuProIlyProGlyIlyAlaThrProArgSerGlnValIleGlnIlyLeuAlaLeu 189
 DB 2654 CAAGTTGCTGCCAAGGTGCAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGT 2713
 QY 190 AlaIlyAlaIleGlnIlyThrLeuGlnIlyAlaIlyThrIlySerAlaIlySerAlaIly 209
 DB 2714 GCTAGTGCAGCTCAAAAGC-----GACTCCCAAGATTAACCAAGCGACTGCAGCTACA 2764
 QY 210 ThrGlnAlaGlnAla---AspGlnThrAsnIlyLeu----- 220
 DB 2765 GCAGAGACGACAGCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2824
 QY 221 GlyLeuGlnIlyGlnAlaIleIlyIlyIlyAspIlyGlnIlyGlnIlyGlnIlyGlnIly 240
 DB 2825 GCGCCAGCTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 2884
 QY 241 AlaAlaGlnIlyLeuSerIlyAspLeuGlnIlyThr----- 252
 DB 2885 AGTGTGCAAGTTCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAG 2944
 QY 253 ---MetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSer 271
 DB 2945 GCCAGTACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3004
 QY 272 IleValAlaAlaIleThrThrCysGlyAlaGlyIlyLeuAlaGlyIlyLeuAlaIleAla 291
 DB 3005 ---GCGGATCCAGCTATCCGAAAGATAGTGGATTCAGTCACTACAGTCAAGGCTGCA 3061
 QY 292 ValGlyAlaAlaAlaAlaGlyIlyAlaAlaGlyIlyAlaAlaAlaAlaAlaAlaAlaThr 311
 DB 3062 AGCGAGGACGCAAAAGCA-----AGCACTAAGCGAGCGCGCAACGACGCGCGG--- 3112
 QY 312 GlnIleThrValGlnAlaValAlaGlnAlaValIlyS---GlnAlaValIleThrAlaVal 330
 DB 3113 -----GCCGTGGTTTCAGTCTGCCAGTCCAGTCCAGTCCAGTCCAGTCCAGTCCAGT 3145
 QY 331 ArgGlnAlaIleThrAlaAlaIleIlyAlaAlaValIlySerGlyIlyIlyAlaIleIle 350
 DB 3146 GAACAGCGCAAGCGCTCAAGTCCGATGTGTGCAAGAGTCCGCGCGCGCGCGCT 3205
 QY 351 LysThrLeuValIlyAlaIleAlaIlyAlaIlyAlaIlySerIlyGlyIlyIlySerIlyVal 368
 DB 3206 AACAGTAATGCGAGTCCCGACGACGATGCGACCAAGCTGTGTGATGACCAACCGCAGCA 3265
 QY 369 -----PheAlaIlyGlyThrGlnIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIly 377


```

Db 233413 TGAGCAGCAGCAAAATCTACTCTATGCAAAAAACAGCAGGAGGAGTAAATCATCGAA 233472
OY 165 yslEuglThrProgluueupProlysProglYval-----ThrProAtgSerg 181
Db 233473 CCTTCGAATTCGCCGAGTTTCACCCGTCGCCGCTTACTGTCCGACCCGGCGGTT 233532
OY 181 luValIleGlunIleGlYleuAla----LeuAlaIleAlaIleGlunIleGlYgluAlaT 200
Db 233533 ACATGTGATGATTTCCGAAGCAATCTGAACCAACCAATGCAAAACCC-----A 233583
OY 200 hrLysSerAlaLeuSerAsnTYraIAserThrGlnAlaGlnAlaAspGlnThrAsnLysL 220
Db 233584 CCAGCAGCCCGCATGATGCTGTTGAAACAACCTCAAGTTCCGAAAAACATCAAC---- 233639
OY 220 eugLYleuGlunIleGlnAlaIleLysIleAspLysGlnArgLunIleGlnIleGlnMetL 240
Db 233640 -----TGGAATCAGGTCGACGCTTGTACGATTAAGGACCTCAACAG----- 233684
OY 240 ysaAlaIleGlnIleLysSerLysAspLeuGlunIYThrMetAspThrValAsnThrValM 260
Db 233685 -----GAGGCATGACACCCCGCAGCAGCAGCAGTGTG 233715
OY 260 etIleAlaValSerValAlaIleThr-----ValIleSerIleValAlaAlaIlePheT 278
Db 233716 TCGTTATCGTCTTAACCGATATGACCTACGCGCGCTGTCGCCCGCGCAGCGCGGAA 233775
OY 278 hrcYsGlyValaGlyLeuAlaIleGlyLeuAlaIleGlyAlaIleValaIleGlyAlaIleAlaIag 298
Db 233776 CGCGCGCGCGCGCAGCGAGGA--GCGGAGAGAGCGCCAGCAGCAAGCGCGGAA 233832
OY 298 lYgLY--AlaIleAlaIleAlaIleAlaIleThrThrValAlaIleThrGlnIleThrValGlnA 317
Db 233833 CTGGAGTACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTACGTGCGAG 233892
OY 317 laValIleGlnAlaValIleLysGlnAlaIleValIleThrAlaValArgGlnAlaIleThrAlaA 337
Db 233893 CTATCCACACACCCCGCAGCAAGCGCAGCTGCGCATCTC-----G 233934
OY 337 laIleLysAlaIleAlaIle-----LysSerGlyIleLysAlaIlePheIleLysT 352
Db 233935 CCAGCCAGACCCGAGTTCTTCTCTATCAACACAAAGAGACATTAACCTACCTGGAAG 233994
OY 352 hrLeuValLys-----AlaIleAlaIleAlaIle 362
Db 233995 AACTGGGCAAAAGCAGACCCGTACAGACGCGCCACCGCCGTAACCGCAGGGGTAC 234054
OY 362 etLYsGlyIleSerLysValAlaIleAlaIleLysGlyThrGlnMetIleAlaLysAsnPhe---- 380
Db 234055 TGCAGGGCTTAAGCGGCTGAACACCCAGCAGCCGACCGCTCAAGCAAAATTTTCACA 234114
OY 381 --ProLysLeuSerLysValIleSerSerLeuThrSerLysTrpValIleThrValGlyValG 400
Db 234115 GTCCCGCAGCAGCAAACTGACCGCTACCTGATCAACAGCACCGCTGCCGCAAGTGTCC 234174
OY 400 lYValIleValAlaIleAlaIleAlaIleGlyLysGlyIleMetGlnMetGlnLeuSerGlnM 420
Db 234175 AATCCGCCATCAAGCGCGCAGCTGTAAGACACACTGGGCGATGCCGACGTGGGTGCGCA 234234
OY 420 etGlnGlnAsnValaIleGlnPheGlnLysGluValGly-----LysLeuGlnAlaA 437
Db 234235 TP-----CTCAGTACCTACACGAGTAGCGAGCAAAATCAAAATTTAATCTCA 234285
OY 437 laAlaAspMetIleSerMetPheThrGlnPheThrGlnIleGlnAlaSerL 457
Db 234286 GCGAAGACTACTT-----GCCACAAGATAGCCCATG 234318
OY 457 ysgLThrGlyIleLysSerAsnGlnMetThrGlnLysAlaIleThr----- 470
Db 234319 CCGTAGCAGGCTGTGATCGGGGTAGCAAAATTAAGGCAAAATGTGGGAGCGCGCAATCG 234378
OY 471 -----LysL 472
Db 234379 GCGCGGCACTCGCGAGATGTGGAGAAACCTGTTGGACGAGCGCATGTAGCGCAAAC 234438

```

```

OY 472 eugLYAlaGlnIleLeuLysAlaTYraIleAlaIleSerGlyAlaIleAlaGlyAla 490
Db 234439 TGTCAACCCAGACAGCCCAAAAGATCATAGCTACTCTGCGAGTTATTCGAGGCAAC 234494

RESULT 28
AAA81490
ID AAA81490 standard; DNA; 1437668 BP.
XX
XX AAA81490;
AC
AC 04-DEC-2000 (first entry)
DT
DE N. meningitidis B full length genome DNA sequence SEQ ID NO:1068.
XX
XX Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;
KM antigen; vaccine; diagnosis; infection; antibacterial; identification;
KM Meningococcus B; Menb; ds.
XX
XX Neisseria meningitidis.
OS
XX WO200022430-A2.
XX
XX 20-APR-2000.
XX
XX 08-OCT-1999; 99WO-US23573.
XX
XX 09-OCT-1998; 98US-0103794.
XX
XX 30-APR-1999; 99US-0132068.
XX
XX (CHIR ) CHIRON CORP.
XX
XX Frazer CM, Hickey E, Peterson J, Tettein H, Venter JC;
XX Masignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;
XX Rappuoli R, Pizza M;
XX WPI: 2000-318079/27.
XX
XX Isolated nucleotide sequences of Neisseria meningitidis which can be
XX used in the diagnosis and treatment of N. meningitidis infection and
XX other Neisseria infections, for example, N.gonorrhoea.
XX
XX Claim 7: Page 866-1272; 1760pp; English.
XX
XX The present invention describes methods of obtaining immunogenic
XX proteins from Neisseria genomic sequences. AAA81453 to AAA82414
XX represent specifically claimed Neisseria meningitidis genomic DNA
XX sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent
XX Neisseria DNA sequences and their corresponding proteins; AAA81254 to
XX AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the
XX isolation of Neisseria meningitidis DNA sequences; and AAA81322 to
XX AAA81452 represent Neisseria meningitidis Menb polynucleotide OMF
XX sequences, which are all used in the exemplification of the present
XX invention. The nucleic acid sequences, protein sequences, and antibodies
XX against them, can be used in the manufacture of a composition. The
XX composition can be used as a medicament (or in the manufacture of a
XX medicament) for treating, preventing or diagnosing infection due to
XX Neisseria bacteria. For example, some of the identified proteins could
XX be components of vaccines against Meningococcus B; against all serotypes;
XX and/or against all pathogenic Neisseriae. Identification of sequences
XX from the bacterium will also facilitate production of biological probes,
XX particularly organism-specific probes. Attempts to make efficacious
XX Meningococcus B vaccines have failed mainly due to antigen tolerance.
XX Multivalent vaccines have also been tried but none have successfully
XX overcome antigenic variability. The provision of further, complete
XX sequences may provide an opportunity to identify secreted or surface
XX exposed proteins that may be presumed targets for the immune system and
XX which are not antigenically variable or at least more conserved than
XX other more variable regions.
XX
XX Sequence 1437668 BP; 344338 A; 353206 C; 385074 G; 355045 T; 5 other;
XX
XX Alignment Scores:

```


Db 1492 AGTGTAACTCTGCACACAGCG---GGAGGAAGTGGCTTCGTGCT 1336

RESULT 26

ID AAA81471/c

AC AAA81471 standard; DNA; 49914 BP.

XX AAA81471:

XX

DT 04-DEC-2000 (first entry)

XX

DE N. meningitidis partial DNA sequence gnm_19 seq ID NO:19.

XX

KW Neisseria meningitidis; Neisseria gonorrhoeae; genome; immunogenic;

XX

KW antigen; vaccine; diagnosis; infection; antibacterial; identification;

XX

KW Meningococcus B; MenB; ds.

XX

OS Neisseria meningitidis.

XX

PN W0200022430-A2.

XX

PD 20-APR-2000.

XX

PF 08-OCT-1999; 99WO-US23573.

XX

PR 09-OCT-1998; 98US-0103794.

XX

PR 30-APR-1999; 99US-0132068.

XX

PA (CHIR) CHIRON CORP.

XX

PI Frazer CM, Hickey E, Peterson J, Tettelin H, Venter JC;

PI Maignani V, Galeotti C, Mora M, Ratti G, Scarselli M, Scarlato V;

PI Rappunli R, Pizsa M;

XX

DR WPI: 2000-318079/27.

XX

PT Isolated nucleotide sequences of Neisseria meningitidis which can be

PT used in the diagnosis and treatment of N. meningitidis infection and

PT other Neisserial infections, for example, N.gonorrhoea .

XX

PS Claim 7: Page 451-466; 1760pp; English.

CC

CC The present invention describes methods of obtaining immunogenic

CC proteins from Neisseria genomic sequences. AAA81453 to AAA82414

CC represent specifically claimed Neisseria meningitidis genomic DNA

CC sequences; AAA81260 to AAA81303 and AAB25620 to AAB25663 represent

CC Neisseria DNA sequences and their corresponding proteins; AAA81254 to

CC AAA81259 and AAA81304 to AAA81321 represent PCR primers used in the

CC isolation of Neisseria meningitidis DNA sequences; and AAA81322 to

CC AAA81352 represent Neisseria meningitidis MenB polynucleotide ORF

CC sequences, which are all used in the exemplification of the present

CC invention. The nucleic acid sequences, protein sequences, and antibodies

CC against them, can be used in the manufacture of a composition. The

CC composition can be used as a medicament (or in the manufacture of a

CC medicament) for treating, preventing or diagnosing infection due to

CC Neisserial bacteria. For example, some of the identified proteins could

CC be components of vaccines against Meningococcus B; against all serotypes;

CC and/or against all pathogenic Neisseriae. Identification of sequences

CC from the bacterium will also facilitate production of biological probes,

CC particularly organism-specific probes. Attempts to make efficacious

CC Meningococcus B vaccines have failed mainly due to antigen tolerance.

CC Multivalent vaccines have also been tried but none have successfully

CC overcome antigenic variability. The provision of further, complete

CC sequences may provide an opportunity to identify secreted or surface

CC exposed proteins that may be presumed targets for the immune system and

CC which are not antigenically variable or at least more conserved than

CC other more variable regions.

XX

XX

Sequence 49914 BP; 13048 A; 11994 C; 11423 G; 13448 T; 1 other;

Alignment Scores:

pred. No.: 0 00807 length: 49914

Score: 182.50 Matches: 123

Percent Similarity: 38.15% Conservative: 83

Query	Match:	21	Mismatches:	219
DB:			Indels:	115
			Gaps:	24
US-09-889-314-2 (1-496) x AAA81471 (1-49914)				
QY	17	AsnIleMetSerGlnValLeuThrSerThrProGlnGlyAlaProGlnGln-----	33	
Db	21316	AAATCGGGGAGCATGCAGACGCCCTACGCCACCGCATTCGATGGCCCCCAAGGCGACATTA	21257	
QY	34	--AspLysLeuSerGlyAsn-GluThr-----LysGlnIleGlnIleThr	47	
Db	21256	ACATCGAAGCGCGGGCGGAAATATGCATCTCATGCCGAGAAAGAGCTCACTACGACAAAC	21197	
QY	47	rArgGlnGlyLysAsnThrGluMetGluSerAspAlaThr-IleValGlyAlaSerGlyL	67	
Db	21196	TTGAAAGCCCAAAAAAGGCGCATTT-----TTCGGCATCAGCTACACCAAGCAC	21146	
QY	67	YsaAspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGlyValAlaAlaG	87	
Db	21145	ACGACACCCACCCACCCAGTCATGTAACC---GGCGTCCCTCAGGAGTGTGA---21093		
QY	87	LYLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyValSerGlyValAlaAla	107	
Db	21092	-----GAATTCGCCATTCGCAATTCAGTTCAGTTGGGTATCCAAATCGCAAGCCACACT	21041	
QY	107	IaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle---GluGluAla	126	
Db	21040	TTGAACACACACATCGGTGGCCGCAAC---ATACCGGAGGCGTATGCGGACGAG	20990	
QY	126	IaSerLysSerMetGluSerThrLeuGluSerLeuGln---SerLeuSerAlaIleGlnM	145	
Db	20989	CACGGCCGATGCCACAGATTATTCCTCCAGAGGATCAAAACACATCCACACAGAAACCG	20930	
QY	145	eLysGlnValGlnAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlySerAlaL	165	
Db	20929	TGACGACGACGAATAATCTACTATGACCAAAAACAGCGCAGGAGCGGCGATACATCGAAA	20870	
QY	165	YsLeuGlnThrProGlnLeuProLysProGlyVal-----ThrProArgSerG	181	
Db	20869	CCTTCATTCGCCGAGTTTACCGGGTCCCGTGGCGCCGACGTCCGACCCGCGGGTT	20810	
QY	181	IuValIleGlnIleGlyLeuAla---LeuAlaLysAlaIleGlnThrLeuGlyValAlaT	200	
Db	20809	ACATTGTGCAATTCGCCAAGGACATCTGAACCCAAATCGAAACCCCTC-----A	20759	
QY	200	hLysSerAlaLeuSerAsnThrLysSerThrGlnAlaGlnAlaAspGlnThrAspLysL	220	
Db	20738	CCAAACACCCGAGTATCTTATTTGAAACAACCTCAAGTTGCGAAAAACATCAAC---	20703	
QY	220	euglyLeuGlnLysGlnAlaIleLysIleAspLysGluArgGlnGluGlyThrGlnIleMetL	240	
Db	20702	-----TGCAATCAGGTGACGCTGTGTTACGATTAATGGACCTCAAAACAG---	20658	
QY	240	YsAlaIleGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValM	260	
Db	20657	-----GAGGCGATGACACCCGACGACGACGACGCTGTG	20627	
QY	260	eLleAlaValSerValAlaIleThr-----ValIleSerIleValAlaAlaIlePheT	278	
Db	20626	TCGTTATGCTGTACCGTATTTGACCTTACGGCGCATGCTGCCCGCGGACGCGCGAA	20567	
QY	278	hrcysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAlaAlaAlaG	298	
Db	20566	CGCGGGCGCGGCGGACGCGCAGA---GCGGAGAGAGCCGACGACGAGACGCGCGGAA	20510	
QY	298	IyGly---AlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla	317	
Db	20509	CTGGAGTGCAGCAGAGAAACGCGACGCCAACACCGGAGTACGACGAGCATTACGTCAG	20450	
QY	317	IaValValGlnAlaValLysGlnAlaValIleThrAlaValAlaValGlnAlaIleThrAlaL	337	
Db	20449	CTATACCAACAGCGCCAGCAACCGCCATCGGCGCAATCTC-----G	20408	

PI Chen Y, Zychlinsky A:
 XX WPI: 1997-393276/36.
 DR P-PSDB: AAW29893.
 XX
 PT Shigella IpaB protein induces apoptosis by binding to
 PT Interleukin-1-beta converting enzyme - useful for treating cancer,
 PT auto-immunity, inflammation, etc. by eradication of unwanted cells
 PS
 XX Claim 3: Page 67-69; 108pp: English.
 XX
 CC This sequence comprises DNA encoding the IpaB protein (see AAW29893)
 CC of Shigella flexneri. This protein induces apoptosis by binding to
 CC Interleukin-1-beta converting enzyme (ICE). A claimed method of
 CC inducing apoptosis in a eukaryotic cell, thereby killing the cell,
 CC comprises: (a) providing to a cell to be killed a DNA molecule (1)
 CC in expressible form which encodes the Shigella IpaB protein or a
 CC functional derivative; and (b) causing (1) to be expressed in the
 CC cell, thereby producing the IpaB protein or derivative which
 CC induces apoptosis and kills the cell. This approach is useful in
 CC treating e.g. cancer, autoimmunity, inflammation and chronic viral
 CC infections. Protein or peptide fragments, and the DNA encoding
 CC them, can also be used to treat diseases that involve apoptotic
 CC mechanisms in their pathogenesis, e.g. AIDS, degenerative diseases
 CC such as Alzheimer's disease, myelodysplastic disorders, ischaemic
 CC injury or toxin-induced liver disease. Methods of gene therapy
 CC relying upon controlled expression of IpaB in a target cell are
 CC disclosed. Also claimed are methods for: inhibiting the
 CC interaction of an apoptosis-inducing protein or peptide with ICE;
 CC detecting a compound capable of inhibiting the binding of IpaB to
 CC ICE; screening a candidate protein for its ability to interact
 CC with IpaB in a cell; and treatment of Shigellosis. Homologues of
 CC IpaB such as yopB of Yersinia enterocolitica and Yersinia
 CC pseudotuberculosis, and sIpb of Salmonella typhimurium and
 CC Salmonella typhi can also be used in the methods of the invention.
 XX
 SQ Sequence 1743 BP: 632 A: 360 C: 302 G: 449 T: 0 other:
 Alignment Scores:
 Pred. No.: 9,91e-05 Length: 1743
 Score: 185.00 Matches: 105
 Percent Similarity: 38.10% Conservative: 84
 Best Local Similarity: 21.17% Mismatches: 171
 Query Match: 7.81% Indels: 136
 DB: 18 Gaps: 21
 US-09-889-314-2 (1-496) x AAT85939 (1-1743)
 QY 42 LysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGluMet----- 55
 DB 316 AAGTCCACGACACGCAAGCAGCAAGCAAAAACCTAGAAATTCGCGATTAATAACACT 375
 QY 55 ----- 55
 DB 376 CTTCATCTGAACCTGAAGACTAACGACACTATGAAAAACMAATTAACTAATAA 435
 QY 56 GluSerAspAlaThrIleAlaGlyAlaSerGlyLysAspLysThrSerSerThrTrpLys 75
 DB 436 AAGCGAGATTCTTAATAAAGACCTAGAAATTAATTAACCAATTAACACAGATTA 495
 QY 76 ThGluThrAlaProGlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLys 95
 DB 496 TCGAACCTCGATCCA-----GAGTCACACGAAAAAGAAAAA 531
 QY 96 AlaGlyAlaAspThrGlyValSerGlyAlaAlaIleThrThraLaserSerThrAlaThr 115
 DB 532 TTAAGCCGGGAGAA-----ATFACAACTCAGCATTAACAAAAAGCGAGAGTT 579
 QY 116 LysIleAlaMetGlnThrSerIleGlnGluAlaSerLysSerMetGluSerThrLeuGln 135
 DB 580 AAA-----GACAGGACATTGATTGACGAGAAAAACCTGCAATTCATAGC----- 624
 QY 136 SerLeuGlnSerLeuSerAlaAlaGlnMetLysGluValGlnAlaValAlaAlaAla 155

DB 625 AAACCTTACAGTAATCAATGCACTCGCAAAAAAGAAATAGACTCT-----TTTCTGCA 678
 QY 156 LeuSerGlyLysSerSerGlySerAlaLysLeuGlnThrProGluLeuProLysProGly 175
 DB 679 TTTTCAAAACACAGCAATCT---GCTGACAGCTATCAACCCAGCAAGAAATTAACCGGA 735
 QY 176 ValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThr 195
 DB 736 CTTCGACGTGTTACTCAATG-----ATGGCACTTTATTCACACTA 777
 QY 196 LeuGlyGluAlaThrLysSerAlaLeuSerAsnTrpAlaSerThrGlnAlaGlnAlaAsp 215
 DB 778 GTTGGAAAAAATATGACAAACATTTAAATAATGATGCTGCTCATTCAGCTCTCCAA 837
 QY 216 GlnThrAsnLysLeuGlyLeuGlnLysGlnAlaIleLysIleAspLysGluArgGlnGlu 235
 DB 838 GAATCAAGAAAAACCTGAATGTGAGAAATCT-----GATGAG 876
 QY 236 TyrGlnGluMetLysAlaIleGlu---GlnLysSerLysAspLeuGlnGlyThrMetAsp 254
 DB 877 TAT-----GCTGCTGAGTACCTTAAAGCAAGAACTCAGACGATATATGGGT 924
 QY 255 ThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAla 274
 DB 925 TGTGTTGGGAAATFACCTGGGCACTT-----TTACATATGTTGTTGTTGCA 975
 QY 275 AlaIlePheThrGlySerGlyAlaGlyLeuAlaGlyLeuAlaGlyAlaValAlaGlyAla 294
 DB 976 GCAGCTTTTCTGAGAGAGCTCTTACGACGTGGACGCTGTGGTACTCTT----- 1029
 QY 295 AlaAlaAlaGlyLysAlaAlaGlyAlaAlaAlaIleThrThrValAlaThrGlnIleThr 314
 DB 1030 -----ANGGTT 1035
 QY 315 ValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValAlaGlnAlaIle 334
 DB 1036 ACGGATCTCTATGACAGCAAGCAGCGCCG----- 1065
 QY 335 ThrAlaIleLysAlaIleValLysSerGlyIleLysAlaPheIleLysThrLeuVal 354
 DB 1066 AATTCCTTCATGACAGCAAGCCCTGAATCCGATCATGAAACAGCTGATGAACCTTAATC 1125
 QY 355 LysAlaIleAlaLysAlaIleLysSerGlyIleSerLysValPheAlaLysGlyThrGln 374
 DB 1126 AAATCCTTCATGACAGCTCTTACAAAA-----ATGCTGAGAGCTTGGGC 1170
 QY 375 MetIleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysTrp 394
 DB 1171 GTGACCTCGAAAAAGCCAAATATGATGGCTCTATTCGGGGGCAATCGACGGGCTCT 1230
 QY 395 ValThrValGlyValGlyValAlaValAlaIleProAlaLeuGlyLysGlyIleMetGln 414
 DB 1231 GTCCTAGTTGACAGCGAGTGTCTCGTAGCC-----ACTGTTGGTAA----- 1272
 QY 415 MetGlnLeuSerGluMetGlnGlnAsnValAlaGlnPheGlnLysGluValGlyLysLeu 434
 DB 1273 -----CAGGACAGCAAAAACCTGCGAGAAATATTTGGCAAAATA 1311
 QY 435 -----GlnAlaAlaAlaAspMetIleSerMetPheThrGln---PheTrpGlnAlaAla 451
 DB 1312 ATAGGTTAAACCCCTACAGACCTTATACCAAGTTTCTCAGAAATTTTCTTCAACTG 1371
 QY 452 SerLysIleAlaSerLysGlnThrGlyLysSerAsnGluMet----- 465
 DB 1372 GACGATTTAATCACTAATGCTGTTGCCAGATTAAATTAATTAATTTCTGTGCGACGGGTGAT 1431
 QY 466 -----ThrGlnLysAlaThrLysLeuGlyAla 474
 DB 1432 GAAGTAATATCCAAACAAATTAATTTCCACCCATTTAAACCAAGCAGTTTATTTAGGAGA 1491
 QY 475 GlnIleLeuLysAlaThrAlaAlaIleSerGlyAlaIleAlaGlyAla 490


```

OY 175 -----GlyValThrProArgSerGluValIleGluIle 185
DB 7696 GCGAACAACAAATGTGGCAAAAAGAAAGCTGCACCCCA---GCAGCAGCAGCAGCAGCA 7640
OY 186 GYLLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThrLysSerAlaLeuSer 205
DB 7639 GCAGCAGCCCGCCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 7583
OY 206 AsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlyLys--- 224
DB 7582 GAGAGCAATGTGACCCCAATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT 7523
OY 225 -----GlnAlaIleLysIleAspLysGluArgGluGlu--- 235
DB 7522 GCCAACAACATCACCAGCAACAGCAGCATTCGACATTCGACATTCGACATTCGACATTCG 7463
OY 236 -----TyrGlnGluMetLysAlaAlaGlnGlnLysSerLysAspLeuGlyThrMet 253
DB 7462 CAGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 7403
OY 254 AspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleVal 273
DB 7402 ACCCAACGCCCAACACA---ATACAAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7346
OY 274 AlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValGly 293
DB 7345 GCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 7286
OY 294 AlaAlaAlaIleGlyAlaAlaGlyAlaAlaAlaAlaIleThrValAlaIleGlnIle 313
DB 7285 TCGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 7226
OY 314 ThrValGlnAlaValAlaGlnAlaValLys 323
DB 7225 ACAGTATCGAATTTGGCACAGTGTCTGCGCA 7196

RESULT 24
AAT85043
ID AAT85043 standard; DNA; 7766 BP.
AC AAT85043;
XX
XX 22-FEB-1998 (first entry)
XX
DE Borrelia variable major protein (VMP)-like sequence (vls) locus.
XX
XX Variable major protein-like sequence; VMP-like sequence; vls locus;
XX vlsE gene; Lyme disease; relapsing fever; therapy; diagnosis;
XX vaccine; ss.
XX
XX OS Borrelia burgdorferi strain B31-5A3 (ATCC 35210).
XX
XX Key Location/Qualifiers
FH misc_feature 205..711
FT /*tag= a
FT /label= vls2
FT /note= "truncated at 5' end"
FT
FT misc_feature 712..1293
FT /*tag= b
FT /label= vls3
FT /label= vls3
FT 1294..1869
FT /*tag= c
FT /label= vls3
FT 1870..2439
FT /*tag= d
FT /label= vls5
FT 2440..3009
FT /*tag= e
FT /label= vls6
FT 3010..3483
FT /*tag= f
FT /label= vls7
FT misc_feature 3484..3990

```

```

FT /*tag= g
FT /label= vls8
FT misc_feature 3991..4548
FT /*tag= h
FT /label= vls9
FT 4549..5058
FT /*tag= i
FT /label= vls10
FT 5059..5652
FT /*tag= j
FT /label= vls11
FT 5653..6219
FT /*tag= k
FT /label= vls12
FT 6220..6789
FT /*tag= l
FT /label= vls13
FT 6846..7373
FT /*tag= m
FT /label= vls14
FT 7274..7946
FT misc_feature 7947..8000
FT /*tag= n
FT /label= vls15
FT /note= "the vls sequence provided is only 7766 bases"
FT
FT misc_feature 7947..8000
FT /*tag= o
FT /label= vls16
FT /note= "the vls sequence provided is only 7766 bases"
FT
FT W09731123-A1.
XX
XX 28-AUG-1997.
XX
XX 20-FEB-1997; 97WO-US02952.
XX
XX 21-FEB-1996; 96DS-0012028.
XX
XX (TEXA ) UNIV TEXAS SYSTEM.
XX
XX PA
XX Barbour AG, Hardham JM, Howell JK, Norris SJ, Weinstein GM;
XX Zhang J;
XX
XX WPI: 1997-435172/40.
XX
XX Nucleic acid encoding variable major protein-like peptide of
XX Borrelia - useful for recombinant production of VMP like protein or
XX peptide, or for diagnosis of Lyme disease
XX
XX Claim 26; Page 101-105; 130pp; English.
XX
XX This DNA sequence comprises the variable major protein (VMP)-like
XX sequence (vls) locus of Borrelia burgdorferi. An infectivity
XX related 28-kb linear plasmid, pBB28la, of B. burgdorferi B31 was
XX isolated by subtractive hybridisation. It contained the vls locus,
XX which consists of 15 silent vls cassettes (vls2-vls16) and an
XX expressed vlsE gene (see AAT85043). Portions of several of the 15
XX silent vls cassette sequences, located approx. 500 bp upstream of
XX vlsE, recombine into the central vlsE cassette region during
XX infection, resulting in antigenic variation of the expressed
XX lipoprotein (see AAW22676) and hence immune evasion. Long-term
XX survival and pathogenesis in the mammalian host. Nucleic acids,
XX including vls2-vls16 sequences, encoding VMP-like proteins, can be
XX used for the recombinant production of VMP-like proteins, or for
XX the diagnosis of Lyme disease (claimed), and may also have
XX application in DNA vaccines.
XX
XX SO Sequence 7766 BP; 2029 A; 867 C; 3011 G; 1858 T; 1 other;
XX
XX Alignment Scores: 0.00049 Length: 7766
XX Pred. No.: 186.00 Matches: 129
XX
XX Score:

```

```
OY 131 GluSerThrLeuGluSerLeuGlnSerLeuSerAlaAlaGlnMetLysGluValGlnAla 150
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2552 CAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2493
OY 151 ValValValAlaAlaAlaLeu-----Ser 157
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2492 ATGCAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2433
OY 158 GlyLysSerSerGlySerAlaLysLeuGlnThrProGluLeuProLysPro----- 174
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2432 GCTGAGCCACGAGGCGCAAGTCAAAAGGTGCAATCCGCAATGACCAAAAGGAGTCC 2373
OY 175 -----GlyValThrProArgSerGluValIleGluIle 185
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2372 GCGCAAAACAATGTTGGCAAAAAGAAAGCTGCACGCCA---GCGACGACGACGACGACG 2316
OY 186 GlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThrLysSerAlaLeuSer 205
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2315 GCAGCAGCGCGCGCGCACGACGACGACGACGACGACGACGACGACGACGACGACG 2259
OY 206 AsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlnLys--- 224
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2258 GAGAGCAATGTGACCCCAATCAATGCGATTAATTAACGCTGCCGCCCAAAAGGCG 2199
OY 225 -----GlnAlaIleLysIleAspLysGluArgGluGlu--- 235
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2198 GCCAAACAATCACCAGACGACGACGACGACGACGACGACGACGACGACGACGACG 2139
OY 236 -----TyrGlnGluMetLysAlaAlaGlnLysSerLysAspLeuGlnGlyThrMet 253
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2138 CAGCGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2079
OY 254 AspThrValaSerThrValMetIleAlaValaSerValaIleThrValIleSerIleVal 273
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2078 ACCAACAAGCCGACCAACA---ATAACAATGTGTGCTGCTTAATGTCTGCTGCTGT 2022
OY 274 AlaAlaIlePheThrCysGlyValaGlyLeuAlaGlyLeuAlaAlaGlyValaValaGly 293
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 2021 GCTGTGTTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1962
OY 294 AlaAlaAlaAlaGlyGlyAlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 313
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 1961 TCGTCACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1902
OY 314 ThrValGlnAlaValaValaGlnAlaValaLys 323
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 1901 ACAGTATCGAATTTGGCACAGTGTGCGA 1872

RESULT 23
ABLO4818/C
ID ABLO4818 standard; cDNA; 18869 BP.
XX
XX ABL04818;
XX
XX 26-MAR-2002 (first entry)
XX
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 8936.
XX
XX Drosophila; developmental biology; cell signalling; insecticide;
XX
XX pharmaceutical; gene; ss.
XX
XX Drosophila melanogaster.
XX
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.
XX
```

```
PA (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW.
XX
XX WPI: 2001-656860/75.
XX
XX P-PSDB: ABB60715.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 8936; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (AB16176-AB130511), expressed DNA
XX sequences (AB101840-AB16175) and the encoded proteins
XX (AB15737-AB12072).
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX SQ Sequence 18869 BP; 5494 A; 3649 C; 3579 G; 6147 T; 0 other;

Alignment Scores:
Pred. No.: 0.00129 Length: 18869
Score: 186.50 Matches: 80
Percent Similarity: 38.00% Conservative: 53
Best Local Similarity: 22.86% Mismatches: 162
Query Match: 7.87% Indels: 55
DB: 23 Gaps: 10

US-09-889-314-2 (1-496) x ABL04818 (1-18869)
OY 25 SerThrProGlnGlyValProGlnGlnAspLysSerGlyAsnGluThrLys----- 42
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 8233 ACACACACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 8177
OY 43 -----GlnIleGlnGlnThrArgGlnGlyLysAsnThr 53
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 8176 CCAGCAGCACCACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 8117
OY 54 GluMetGluSerAspAlaThrIleAlaGlyAlaSerGlyLysAspLysThrSerThr 73
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 8116 GTTGCCGAGGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 8057
OY 74 -----ThrLysThrGluThrLysThrLysThrLysThrLysThrLysThrLysThr 90
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 8056 GCCATGCCACGCGACCAACCAACCAACCAACCAACCAACCAACCAACCAACCAACCA 7997
OY 91 SerGluSerGlnLysAlaGlyAlaAspThrGlyValSerGlyAlaAlaAlaThrThrAla 110
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 7996 AATTGAACGAAACGACGAGGCGCAAAATTAATATGCGGACCGGCAACCTGAA 7937
OY 111 SerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGluGlnAlaSerLysSerMet 130
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 7936 CAACAACGCGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAA 7877
OY 131 GluSerThrLeuGluSerLeuSerLeuSerAlaAlaGlnMetLysGluValGlnAla 150
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 7876 CAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 7817
OY 151 ValValValAlaAlaAlaLeu-----Ser 157
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 7816 ATGCAACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 7757
OY 158 GlyLysSerSerGlySerAlaLysLeuGlnThrProGluLeuProLysPro----- 174
   :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::
Db 7756 GCTGAGCCACGAGGCGCAAGTCAAAAGGTGCAATCCGCAATGACCAAAAGGAGTCC 7697
```

```

Db 2980 GTTCCGAGCGAGAAACATCTACCGAAGGGTTGAACATCTGGCAACGCCGTA 3039
QY 74 -----ThrlsThrglnThrAlaProGlnGlnGlyValAlaIlaGlyLysGlnSer 90
Db 3040 GCCATGCCAACGGCCACACCAACCAACCAACCAACCAACCAACCAACCAATGA 3099
QY 91 SerGlnSerGlnLysAlaGlyAlaIlaSerThrGlyValSerGlyAlaIlaThrAla 110
Db 3100 AAATTGAACGACGACGAGCCAAAATAATAGCGGCGGCGGCGGCGGCGGCGGCGG 3159
QY 111 SerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlnGlnLysSerMet 130
Db 3160 CAACAACAGGCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3219
QY 131 GlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaIlaGlnMetLysGlnValAla 150
Db 3220 CACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 3279
QY 151 ValValValAlaIlaIleu-----Ser 157
Db 3280 ATGCACACGCGACGACATTTGGTCCACAAACGAAACGAAACGAAATTTGACGGAAGCC 3339
QY 158 GlyLysSerSerGlySerAlaLysLeuGlnThrProGlnLeuProLysPro----- 174
Db 3340 GCTGAGCGCAACGGCGCAAGTCAAAAGTGCATACCGCAATTCGACAAACGAAAGTCC 3399
QY 175 -----GlyValThrProThrSerGlnValIleGlnIle 185
Db 3400 GGCAAAACAAATGTTGGCAAAACAAAGCTGCACGCCA---GCAGCAGCAGCAGCAGCA 3456
QY 186 GlnLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnLysAlaThrLysSerAlaLeuSer 205
Db 3457 GCAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3513
QY 206 AsnThrAlaSerThrGlnAlaGlnAlaSerGlnThrAsnLysLysGlnLysLys--- 224
Db 3514 GAGAGCAATGTGACCCCAATCATTCATTAACGCTGGCGCGCGCGCGCGCGCGCGCG 3573
QY 225 -----GlnAlaIleLysIleAspLysGlnLysGlnLysGlnLys--- 235
Db 3574 GCCAAACATCATCCGACAGCAGCAGCATTTGCAATTTGCAACAGCAGCAGCAGCAG 3633
QY 236 -----TyrGlnGlnMetLysAlaIlaGlnLysSerLysAspLeuGlnGlyThrMet 253
Db 3634 CAGCGCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3693
QY 254 AspThrValAsnThrValMetIleAlaIlaSerValAlaIlaIleThrValIleSerIleVal 273
Db 3694 ACCACACGCGCCACACAACA---ATAACAATTTGTTGCTGCTGCTGCTGCTGCTGCT 3750
QY 274 AlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaIlaGlyAlaIlaValGly 293
Db 3751 GCGTTGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3810
QY 294 AlaAlaIlaIlaGlyGlyAlaIlaGlyAlaIlaIlaIlaIlaIlaIlaIlaIlaIla 313
Db 3811 TCGCTCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3870
QY 314 ThrValGlnAlaValValGlnAlaValLys 323
Db 3871 ACAGTATCGAATTTGGCAGCAGTCTCGCA 3900

```

RESULT 22
ABLI7520/c
ID ABLI7520 standard; DNA; 6771 BP.

XX ABLI7520;
XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster genomic polynucleotide SEQ ID NO 4033.
XX

```

KW Drosophila; developmental biology; cell signalling; insecticide;
KM pharmaceutical; gene; ds.
XX Drosophila melanogaster.
OS WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US09231.
XX
XX 23-MAR-2000; 2000US-191637P.
XX
XX 11-JUL-2000; 2000US-0614150.
XX
XX (PEKE ) PE CORP NY.
XX
XX Venter JC, Adams M, Li PWD, Myers EW;
XX WPI; 2001-656860/75.
XX
XX New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signalling and cell-cell
XX interactions -
XX
XX Claim 1; SEQ ID NO 4033; 21pp + Sequence Listing; English.
XX
XX The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA
XX sequences (ABLI01840-ABLI6175) and the encoded proteins
XX (ABBS7737-ABBS7072)
XX
XX The sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX
XX Sequence 6771 BP; 1504 A; 1470 C; 1519 G; 2278 T; 0 other:
XX
XX
XX Alignment Scores:
XX Pred. No.: 0.000382 Length: 6771
XX Score: 186.50 Matches: 80
XX Percent Similarity: 38.00% Conservative: 53
XX Best Local Similarity: 22.86% Mismatches: 162
XX Query Match: 7.87% Indels: 55
XX DB: 23 Gaps: 10
XX
XX
XX US-09-889-314-2 (1-496) x ABLI7520 (1-6771)
XX
XX 25 SerThrProGlnGlnValProGlnGlnAspLysLeuSerGlnLysGlnLys----- 42
XX ::::::::::::::::::::
XX Db 2909 ACACACACCAAAACACACCAAGTT---AAGCACACACACAGTGAAGTGCATCA 2853
XX
XX 43 -----GlnIleGlnGlnThrAspGlnGlnLysAsnThr 53
XX ::::::::::::::::::::
XX Db 2852 CCAGCAGCACAAGACACAATGCGCAAGTCAAGTCAACAAAGGCGAGCAACACT 2793
XX
XX 54 GlnMetGlnSerAspAlaThrIleAlaGlyAlaSerGlyLysAspLysThrSerSerThr 73
XX ::::::::::::::::::::
XX Db 2792 GTTGGCAGGCGCAAAACCAATCTCACCAAGGGTTGAACAATCTGGCAAGCCGTA 2733
XX
XX 74 -----ThrlsThrglnThrAlaProGlnGlnGlyValAlaIlaGlyLysGlnSer 90
XX ::::::::::::::::::::
XX Db 2732 GCCATGCCAACGCGCAACCAACCAACCAACCAACCAACCAACCAACCAATGA 2673
XX
XX 91 SerGlnSerGlnLysAlaGlyAlaIlaSerThrGlyValSerGlyAlaIlaIlaIlaIla 110
XX ::::::::::::::::::::
XX Db 2672 AAATTGAACGACGACGAGCCAAAATAATAGCGGCGGCGGCGGCGGCGGCGG 2613
XX
XX 111 SerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlnGlnLysSerMet 130
XX ::::::::::::::::::::
XX Db 2612 CAACAACAGGCAACGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2553
XX

```



```

OY 130 tcluserthrleugluserleuglnserleuseralaalaglnmetlysglvalglua1 150
DB 6814 ACCAGTGNACA-----GCACTGATAGACAGACAGCATAGTGTGCACG 6858
OY 150 aValValValAla---AlaLeuSerGlyLysSerSerGlySerAlaLysLeuGlu---- 167
DB 6859 AGTGATAGCAGCAGACAGTGTAGCAGCACTGATAGCAGCAGCAAGTGTAGTGA 6918
OY 168 -----ThrProGluLeuProLysProGlyValThrProArgSerGluVa 182
DB 6919 ACCAGTAATAGTAGTGACAAACAGCAATAGCACTAGACAGCAGACAGCACT 6978
OY 182 lllleuglilegylleualaleualalysalaalileglnthrleugliala1thrlysse 202
DB 6979 GATAGCAGCAGACAGCAGTATATAGTGTGCACG -AGCAATAGCGGTGACAGCAGCAG 7037
OY 202 rAlaLeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLe 222
DB 7038 CAGTGACAGCAGTGTATAGCAATAGCAGCAGCAGCACTAGCAGCAGCAGCATAG 7097
OY 222 uGluLysGlnAlaIleLysIleAspLysGluArgGluGluTyrGlnGlnGlnmetLysAla1 242
DB 7098 CAGTGACAGCAGTGTATAGCAGCAGCAGCAGCAGTGTATAGCAGCAGCAGTGA 7157
OY 242 aGluGlnLysSerLysAspLeuGlnGluThrMetAspThr-ValAsn----- 257
DB 7158 TAGCAGTGTACAGCAGCAGCAGCAGTGTATAGTGTAGTGTAGCAGCAGCAGCAGCAG 7217
OY 258 --ThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIle- 276
DB 7218 TGACAGCAGCAGTGTATAGCAGCAGCAGTGTAGCAGCAGCAGCAGCAGCAGCAATAG 7277
OY 277 -----PheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaVala1G 293
DB 7278 CAGTGACAGCAGTGTATAGCAGCAGCAGCAGCAGTGTATAGCAGCAGCAGTGA 7337
OY 293 lYAlaAlaAlaAlaGlyGlnAlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAla1 313
DB 7338 CAGCAGTGTATAGCAGCAGCAGCAGCAGTGTATAGCAGCAGTGTATAGCAGCAGCAGCAG 7397
OY 313 lEthrValGlnAlaValValGlnAlaValLysGlnAlaValIleThrAlaVala1ArgGlna 333
DB 7398 TGACA-----GCAGTGACAGCAGTGTAAAGCAGCAGCAGCAGCAGTATAGCAGCAG 7448
OY 333 lAlleThrAlaAlaIleLysAlaAlaValLysSerGlyLysAlaIleThrL 353
DB 7449 CAGTGACAGCAGCAG-----ACAGCAGTGTACAGCAGCAGTATAGCAGCAGCAGCAGCAG 7502
OY 353 euValLysAlaIleAlaLysAlaIleSerLysGlyLysSerLysValPheAlaLysGlyT 373
DB 7503 CAGCGATAGCAGTGTATAGCAGCAGCAATAGCA-----GTGATAGCA---GCCA 7544
OY 373 hGlnMetIleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerL 393
DB 7545 CAGCACTGTATAGCA-----GTGACAGCAGCAGCAGCAGCAG 7577
OY 393 yStrIValThrValGlyValGlyValValAlaAlaAlaProAlaLeuGlyLysGlyIleM 413
DB 7578 CCATAGCAGCAGCAGCAGTGTATAGTGTATAGCAGCAGCAGTGTATAGCAGCAGCAGCAG 7637
OY 413 eGlnMetGlnLeuSerGlnMetGlnGlnAlaAsnValAlaGlnPheGlnLysGlnVala1Gly 433
DB 7638 CAGTGACAGCAGCAGCAGCAGTGTATAGCAGCAGCAGCAGCAGTGTATAGCAGCAGCAGCAG 7697
OY 433 yslLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPheThrGlnAlaSerL 453
DB 7698 CAGCAGTGTACAGCAGCAGTGTATAGCAGCAGCAGCAGCAGCAGTGTATAGCAGCAGCAGCAG 7756
OY 453 yslIleAlaSerLysGlnThrGlyGlnSerAsnGlnMetThrGlnLysAlaThrLysLeuG 473
DB 7757 GGTATAGCAGTGTATAGCAGCAGTGTATAGCAGCAGTGTATAGCAGCAGCAGCAGCAGTGT 7816

```

```

OY 473 ly 473
DB 7817 GT 7818

RESULT 21
ABLI7521
ID ABLI7521 standard; DNA; 4638 BP.
XX
AC ABLI7521;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 4036.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN W0200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
XX
PR 11-JUL-2000; 2000US-0614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PMD, Myers EW;
XX
DR WPI; 2001-656860/75.

PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1: SEQ ID NO 4036; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABLI6176-ABLI30511), expressed DNA
CC sequences (ABLI01840-ABLI6175) and the encoded proteins
CC (ABBI7737-ABBI2072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pt_sequences.
XX
SQ Sequence 4638 BP; 1594 A; 1190 C; 1087 G; 767 T; 0 other:

Alignment Scores:
Pred. No.: 0.000244 Length: 4638
Score: 186.50 Matches: 80
Percent Similarity: 38.00% Conservative: 53
Best Local Similarity: 22.86% Mismatches: 162
Query Match: 7.87% Indels: 55
DB: 23 Gaps: 10

US-09-889-314-2 (1-496) x ABLI7521 (1-4638)

OY 25 SerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLys----- 42
DB 2863 ACAACACCAACAAACCAACCAAGTT---AAGGCAGCAACAGATGATGCAAGTCAACA 2919
OY 43 -----GlnIleGlnGlnThrArgGlnGlnLysAsnThr 53
DB 2920 CCAGCAGCAGCAACAAACAAACATGCGCAAGATTCAGCATACAAATGCGCAGCAACACT 2979
OY 54 GluMetGluSerAspAlaThrIleAlaGlyAlaSerGlyLysAspLysThrSerSerThr 73

```



```

XX MPI: 2000-679515/66.
DR P-PSDB: AAB19772.
XX
PT Detecting small integrin-binding ligand N-linked glycoproteins for
PT detection of a tumor or protection against a complement mediated immune
PT response, comprises detection where Factor H is not an inhibitor -
XX
PS Disclosure: Page 98-100; 110pp; English.
XX
CC The present sequence is that of DNA encoding human dentin
CC sialoprophosphoprotein (DSPP, see AAB19772), a member of the small
CC integrin binding ligand, N-linked glycoproteins (SIBLINGS) family.
CC The invention provides methods and compositions for exploiting the
CC discovery that members of the SIBLINGS family bind to complement
CC Factor H, conferring resistance to complement mediated lysis. A
CC claimed method of conferring protection against a complement
CC mediated immune response involves providing a reservoir or other
CC supply in the subject's body so that a SIBLINGS protein can be
CC dispersed to interfere with complement mediated lysis and
CC inflammation. This protects cells that are grafted onto foreign
CC tissue or bone marrow cells introduced into a foreign host. The
CC SIBLINGS protein can be BSP, OPN, DMP1 or DSPP. A method of
CC detecting a SIBLINGS protein in a sample from a subject suspected
CC of having abnormal bone turnover, especially osteoporosis, is also
CC claimed.
XX
SQ Sequence 8201 BP; 2801 A; 1477 C; 1873 G; 2050 T; 0 other:
XX
Alignment Scores:
Pred. No.: 0.0004 Length: 8201
Score: 187.00 Matches: 106
Percent Similarity: 39.24% Conservative: 91
Best Local Similarity: 21.12% Mismatches: 251
Query Match: 7.89% Indels: 56
DB: 21 Gaps: 11
XX
US-09-889-314-2 (1-496) x AAA88864 (1-8201)
OY 1 AASPThAsmMetSerIleSerSerSerGlyProAspAsnGlnIlyAsnIleMetSer 20
DB 6394 GACAGTAGAGGATAGAGGATAGAGGATAGAGGATAGAGGATAGAGGATAGAGG 6453
OY 21 GlnValIleThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlu 40
DB 6454 AGTGATAGTAGTACAGCAGCAGCAGCAGCAGTATAGCAGCAGCAGCAGCAGTATAG 6513
OY 41 ThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThr 60
DB 6514 AGCAGTAGATAGCAGCAGCAGCAGCAGTATAGCAGCAGCAGCAGTATAGCAGTATAG 6573
OY 61 IleAla-----GlnAlaSerIlyLysAspLysThr 70
DB 6574 GACAGCAGTAGATAGCAGCAGCAGCAGCAGTATAGCAGCAGCAGCAGCAGTATAGTATAG 6633
OY 71 SerSerThrThrLysThrGlnThrAlaProGlnGlnGlyValAlaIleGlnLysGluSer 90
DB 6634 GATAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 6693
OY 91 SerGluSerGlnLysAlaGlyAlaAspThrGly-ValSerGlyAlaAlaAlaThrThrAl 110
DB 6694 GACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 6753
OY 110 AsSerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlnGlnAlaSerLysSerIle 130
DB 6754 AGCAACAGCAGCAGCAGCAGCAGTATAGTATAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 6813
OY 130 tGluSerThrLeuGlnLysLeuGlnSerLeuSerAlaIleGlnMetLysGlnValGlnAl 150
DB 6814 AGCAGTAGACA-----GCAGTAGATAGCAGCAGCAGCAGCAGTATAGTATAGCAGCAG 6858
OY 150 aValValValAlaAla--AlaLeuSerGlyLysSerSerGlySerAlaLysLeu----- 167
DB 167

```

```

DB 6859 AGTAGTAGCAGCAGCAGCAGTAGTATAGCAGCAGCAGCAGCAGTATAGCAGTGA 6918
OY 168 -----ThrProGluLeuProLysProGlyValAlaThrProArgSerGluVal 182
DB 6919 AGCAGTATAGTAGTATAGCAGCAGCAGCAGCAGTATAGCAGCAGCAGCAGTATAG 6978
OY 182 IleGlnIleGlyLeuAlaIleValAlaLysAlaIleGlnThrLeuGlnGlyValAlaThrLysSe 202
DB 6979 GATAGCAGCAGCAGCAGTATAGTATAGTATAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7037
OY 202 rAlaLeuSerAsnTyrrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLe 222
DB 7038 CAGTAGCAGCAGCAGTATAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7097
OY 222 uGlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGln 242
DB 7098 CAGTAGCAGCAGCAGTATAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7157
OY 242 aGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThr-ValAsn----- 257
DB 7158 TAGCAGTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTATAGTATAGCAGCAGCAGCAG 7217
OY 258 --ThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaIle- 276
DB 7218 TGACAGCAGCAGCAGTATAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7277
OY 277 -----PheThrGlyGlnAlaGlyLeuAlaGlyLeuAlaIleGlyAlaIleAlaValAlG 293
DB 7278 CAGTAGCAGCAGCAGTATAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7337
OY 293 lYAlaAlaIleAlaGlyGlyAlaIleAlaIleAlaAlaAlaAlaAlaAlaAlaAlaAlaThrValAlaThrGlnI 313
DB 7338 CAGCAGTAGTATAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7397
OY 313 lThrValAlaIleAlaValAlaIleAlaValAlaIleGlnAlaValIleThrAlaValAlaArgGlnA 333
DB 7398 TGACA-----GCAGTAGCAGCAGTATAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7448
OY 333 lAlaIleThrAlaIleAlaLysAlaIleAlaValLysSerGlyLysAlaPheIleLysThrL 353
DB 7449 CAGTAGCAGCAGCAG-----ACAGCAGTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7502
OY 353 euValLysAlaIleAlaLysAlaIleSerLysGlyLysSerLysValPheAlaLysGlyLT 373
DB 7503 CAGCAGTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7544
OY 373 hGlnMetIleAlaLysAsnPheProLysLeuSerLysValIleSerSerLysLeuThrSerL 393
DB 7545 CAGCAGTAGTATAGCA-----GTGACAGCAGCAGCAGCAGCAG 7577
OY 393 yStrpValThrValGlyValAlaGlyValAlaValAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 413
DB 7578 CGATAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7637
OY 413 eGlnMetGlnLeuSerGlnMetGlnGlnGlnAlaValAlaAlaGlnPheGlnLysGlnValAlaGlyL 433
DB 7638 CAGTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7697
OY 433 yLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPheThrGlnAlaAlaSerL 453
DB 7698 CAGCAGTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7756
OY 453 yIleAlaSerLysGlnThrGlyGlnSerAsnGlnLysAlaThrGlnLysAlaThrLysLeuG 473
DB 7757 GTGATAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 7816
OY 473 lY 473
DB 7817 GT 7818
RESULT 20
AB073537
ID AB073537 standard; DNA; 8201 BP.

```

CC library which contain sequences that hybridise to the target sequence and
 CC isolating the nucleic acid molecules from the members; or (b) isolating
 CC mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid
 CC molecules whose nucleotide sequence is homologous to amplification
 CC primers derived from the fragment of the *S. pneumoniae* genome to prime
 CC the amplification and isolating the amplified sequences. The computer
 CC readable medium can be used in a computer-based system for identifying
 CC fragments of the *S. pneumoniae* genome of commercial importance, or
 CC expression modulating fragments of the *S. pneumoniae* genome. Products
 CC from the present invention can be used in diagnosis kits and assays, and
 CC pharmaceutical compositions and vaccines for *S. pneumoniae*.

XX Sequence 1436 BP; 357 A; 465 C; 331 G; 282 T; 1 other;

Alignment Scores:

Pred. No.:	4,71e-05	Length:	1436
Score:	188.00	Matches:	79
Percent Similarity:	37.30%	Conservative:	59
Best Local Similarity:	21.35%	Mismatches:	214
Query Match:	7.94%	Indels:	18
DB:	19	Gaps:	6

US-09-889-314-2 (1-496) x AAV52498 (1-1436)

QY 2 ThrAsnMetSerIleSerSerSerSergIlyProAspAsnGlnLysAsnIleMetSergIn 21
 DB 298 ACAAGTGGCTGCGCTTCAGCAAGTACCTCAGCTGTAATCAGCATCAACGATGCAAC 357
 QY 22 ValIeuThrSerThrProGlnGlyValProGlnAsnLysSergIlyAsnGluThr 41
 DB 358 GCTTCAGCATCAACA-----AGTCCTCAGCTTCAGCAAGTATCTCAGCTCGAATCG 411
 QY 42 LysGlnIleGlnGlnThrArgGlnGlnLysAsnThrGluMetGluSerAspAlaThrIle 61
 DB 412 GCATCAACGAGTGGCTGCGCTTCAGCAAGTACCTCAGCTGTAATCAGCATCAACGATG 471
 QY 62 AlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGln 81
 DB 472 GCTTCGCTTCAGCTCAACGAGTGGCTGTAATCAGCATCAACGATGCAACGATGCAAC 531
 QY 82 GlnGlyValAlaAlaGlyLysGluSerSergIlyLysGlnLysAlaGlyAlaAspThrGly 101
 DB 532 GCAACGACATCAGCTTCGTAATCTGCATCAACGATGCGTCAAGCTCAGCATCGACAAGC 591
 QY 102 ValSergIlyAlaAlaAlaThrThrAlaSerAsnThrAla----- 114
 DB 592 GCCTCAGCTTCAGCAAGTACCTCAGCTGTAATCTGCATCAACGATGCGTCAAGCTC 651
 QY 115 -----ThrLysIleAlaMetGlnThrSerIle-----GluGluAlaSerLysSerMetGlu 131
 DB 652 GCGTCGACAAAGTGGCTGCGCTTCAGCAAGTACCTCAGCTGTAATCTGCATCAACGATG 711
 QY 132 SerThrLeuGlnSerLeuGlnSerLeuSerAlaAlaGlnMetLysGluValGluAlaVal 151
 DB 712 GCCTCAGCAAGTACCTCAGCTGTAATCTGCATCAACGATGCGTCAAGCTCAGCAAGT 771
 QY 152 ValValAlaAlaLeuSerGlyLysSerSergIlyLysLeuGlnThrProGluLeu 171
 DB 772 ACTAGTCATCAGCTTCAGCAAGTACCTCAGCTGTAATCTGCATCAACGATGCGTCAAGT 831
 QY 172 ProLysProGlyValThrProArgSergIlyValIleGluIleGlyLeuAlaLeuAlaLys 191
 DB 832 GCTTCAGCAAGTACCTCAGCTGTAATCTGCATCAACGATGCGTCAAGCTCAGCAAGT 885
 QY 192 AlaIleGlnThrLeuGlnGlyAlaThrLysSerAlaLeuSerAsnThrLysSerThrGln 211
 DB 886 GCAAGTACCTCAGCTGTAATCTGCATCAACGATGCGTCAAGCTCAGCAAGTACCAAGT 945
 QY 212 AlaGlnAlaAspGlnThrAsnLysLeuGlu-----LysGlnAlaIleLys 228
 DB 946 GCATCAGCTTCAGCAAGTACCTCAGCTGTAATCTGCATCAACGATGCGTCAAGTACCAAGT 1005
 QY 229 IleAspLysGluArgGlnGluThrGlnGluMetLysAlaAlaGlnGlnLysSerLysAsp 248

DB 1006 GCATCAGCAAGTACCTCAGCTGTAATCTGCATCAACGATGCGTCAAGTACCAAGT 1065
 QY 249 LeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaValSerValAlaThr 268
 DB 1066 GCGTCGCTTCAGCAAGTACCTCAGCTGTAATCTGCATCAACGATGCGTCAAGTACCAAGT 1125
 QY 269 ValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAla 288
 DB 1126 GCATCAACGAGTGGCTGCGCTTCAGCAAGTACCTCAGCTGTAATCTGCATCAACGATG 1185
 QY 289 GlyAlaAlaValAlaGlyAlaAlaAlaGlyGlyAlaAlaGlyAlaAlaAlaAlaThr 308
 DB 1186 GCGTCAGCTTCGCAATCAACGATGCGTCAAGTACCAAGTACCAAGTACCAAGTACCAAGT 1245
 QY 309 ValAlaThrGlnIleThrValAlaGlnAlaValAlaGlnAlaValLysGlnAlaValIleThr 328
 DB 1246 GCATCAACGAGTGGCTGCGCTTCAGCAAGTACCTCAGCTGTAATCTGCATCAACGATG 1305
 QY 329 AlaValArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSergIlyLeuLysAla 348
 DB 1306 GCGTCGCTTCAGCAAGTACCTCAGCTGTAATCTGCATCAACGATGCGTCAAGTACCAAGT 1362
 QY 349 PheIleLysThrLeuValLysAlaIleAla 358
 DB 1363 TCAGCATCAACGATGCGTTCGCTGCGCA 1392
 RESULT 19
 AAA88864
 ID AAA88864 standard; DNA: 8201 BP.
 AC AAA88864;
 XX 19-FEB-2001 (first entry)
 DT Human dentin sialoprophosphoprotein gene.
 XX
 DE Human dentin sialoprophosphoprotein; DSPP; human; SIBLINGS;
 KW Integrin-binding ligand; diagnosis; antiinflammatory; therapy;
 KW osteoporosis; chromosome 4; ds.
 XX
 OS Homo sapiens.
 XX
 FT Key Location/Qualifiers
 FT CDS 2387..7896
 FT /*tag= a
 FT /note= "contains exons"
 FT exon 2387..2437
 FT /*tag= b
 FT Intron 2438..3576
 FT /*tag= c
 FT exon 3577..3660
 FT /*tag= d
 FT Intron 3661..3793
 FT /*tag= e
 FT exon 3794..4780
 FT /*tag= f
 FT Intron 4781..5256
 FT /*tag= g
 FT exon 5257..7896
 FT /*tag= h
 PN WO200062065-A1.
 PD 19-OCT-2000.
 PF 07-APR-2000; 2000WO-US09349.
 PR 09-APR-1999; 99US-0128468.
 XX (USSH) US DEPT HEALTH & HUMAN SERVICES.
 XX Fisher LW, Fedarko NS, Young MF;
 PI


```

Db 1249 ----TCGCGTTACAGCAGTACAGCCGCTCAGCAGCTTCCGCTTCA 1305
OY 358 AAlaValAlaIleSerIysGlyIleSerIysValAlaIleSerIysGlyIleAla 377
Db 1306 GCG-----TCAACGAGTGGCTGTGAGTCAGCATCAACGAGTACGCTCAGCA 1356
OY 378 LysAsnProIleSerIysValIleSerIysLeuThrSerIysTrpValThrVal 397
Db 1357 AGCAGCATCAGCTTGTGATGTGATCAGCAGCGCTCAGCAGCTCAGCAGCAGC 1416
OY 398 GlyValGlyValValAlaAlaAlaProAlaLeuGlyLys-GlyIleMetGlnIle 417
Db 1417 TCAGCTTACAGCAGTACAGTACGCTCAGCAGCTCAGCAGTACGCTCAGCAG 1476
OY 417 userGlnMetGln-----GlnAsnValAlaGlnPh 427
Db 1477 TCGACAGTGGCTGGCCGCTCAGCAGCTGATCTGATCGGCATCAACAGTGGCTCAGCT 1536
OY 427 eGlnLys----- 429
Db 1537 CAGCAGTACTAGTGCATCAGCTTACAGCATCAACGAGTGCATCGGCTTCCGCTCAACCA 1596
OY 430 -----GluValGlyIysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPh 447
Db 1597 GTGCATCAGAGTCAGCAAGTACAGTGCCTACTTCC----- 1633
OY 447 eTrpGlnGlnAlaSerIysIleAlaSerIysGlnThrGlyIleUserAsnGlnMetThrGln 467
Db 1634 ----GCATCAACAGAGTCCCTGCTTACAGCAGCAGTACGATCTGATCAGGCTCAGC 1689
OY 467 nLysAlaThrLysLeuGlnAlaGlnIleLeuLysAlaThrAlaAlaIleSerGlyAlaIle 487
Db 1690 CAGTGGCTTCCGCTTACAGCAGTACAGTGCCTTACGCTTACGATCAACAGCAGCGCTCGGC 1749
OY 487 eAlaGlyAla 490
Db 1750 CTCAGCAAGC 1759

RESULT 17
ID AAV03553 standard; DNA: 15620 BP.
XX
AC AAV03553;
XX
DT 22-OCT-1998 (first entry)
XX
DE DNA sequence that is specific for Neisseria meningitidis.
XX
KW N. gonorrhoeae; N. lactamica; chromosome 22491; region 1; region 2;
KW region 3; pathogenicity; blood-brain barrier; diagnosis; infection;
KW meningitis; ss.
XX
OS Neisseria meningitidis.
XX
FH Key Location/Qualifiers
FT CDS 1330..2973
FT FT /*tag= a
FT FT /*tag= "encodes AAW42633"
FT FT /*tag= b
FT FT /*tag= "encodes AAW42634"
FT FT 9044..9475
FT FT /*tag= c
FT FT /*tag= "encodes AAW42635"
FT FT 10127..12121
FT FT /*tag= d
FT FT /*tag= "encodes AAW42636"
FT FT 12118..12609
FT FT /*tag= e
FT FT /*tag= "encodes AAW42637"
FT FT 12794..13056
FT FT /*tag= f
FT CDS

```

```

FT FT /note= "encodes AAW42639"
FT FT 13297..14238
FT FT /*tag= g
FT FT /*tag= "encodes AAW42640"
FT FT 14241..15176
FT FT /*tag= h
FT FT /note= "encodes AAW42641"
FT FT
FT FT W09802547-A2.
FT FT
FT FT 22-JAN-1998.
FT FT
FT FT 11-JUL-1997; 97WO-FR01295.
FT FT
FT FT 12-JUL-1996; 96FR-0008768.
FT FT
FT FT (INRM ) INSERM INST NAT SANTE & RECH MEDICALE.
FT FT (PLAC ) MAX PLANCK GES FOERDERUNG WISSENSCHAFTEN.
FT FT (SMK ) SMITHKLINE BECHAM.
FT FT
FT FT Nassif X, Tinsley C, Achman M, Merker P, Ruelle J;
FT FT Vinals C;
FT FT
FT FT WPI: 1998-110594/10.
FT FT P-PSDB: AAW42633, AAW42634, AAW42635, AAW42636, AAW42637, AAW42638,
FT FT AAW42639; P-PSDB: AAW42640, AAW42641.
FT FT
FT FT Genes present in Neisseria meningitidis but not other Neisseria
FT FT species - and related host cells, RNA, anti-sense sequences,
FT FT polypeptide(s) and antibodies, useful for diagnosing Neisseria
FT FT meningitidis infection and in protective vaccines
FT FT
FT FT Claim 8; Pages 67-78; 150pp; French.
FT FT
FT FT AAV03518-53 represent sequences that are present in Neisseria
FT FT meningitidis but not in N. gonorrhoeae or N. lactamica, except for the
FT FT CC genes involved in biosynthesis of the capsule polysaccharide, trpA or C,
FT FT CC opc, porA, rotamase, sequence IC1106, Iga protease, pilin, pilC,
FT FT CC proteins which bind transferrin and opacity proteins. The sequences are
FT FT CC found on chromosome 22491, mainly (or within 20 kb) between tufa and pili
FT FT CC (region 1), pilO and lambda-740 (region 2) or argF and opaB (region 3).
FT FT CC The DNA sequences are responsible for the differences in pathogenicity
FT FT CC between N. meningitidis and N. gonorrhoeae, specifically they include the
FT FT CC genes that allow N. meningitidis to cross the blood-brain barrier. DNA
FT FT CC sequences common to N. meningitidis and N. gonorrhoeae, but absent from
FT FT CC N. lactamica, are responsible for colonisation and penetration of the
FT FT CC mucosa. The DNA sequences can be used to produce probes and primers, and
FT FT CC antibodies produced against the encoded proteins are used in standard
FT FT CC hybridisation/immunoassay processes for diagnosis of N. meningitidis
FT FT infection, particularly meningitis.
FT FT
FT FT Sequence 15620 BP; 4806 A; 3505 C; 3556 G; 3753 T; 0 other;
FT FT
FT FT Alignment Scores:
FT FT Pred. No.: 0 000437 length: 15620
FT FT Score: 191.50 Matches: 124
FT FT Percent Similarity: 38.33% Conservative: 83
FT FT Best Local Similarity: 22.96% Mismatches: 218
FT FT Query Match: 8.08% Indels: 115
FT FT DB: 19 Gaps: 24
FT FT
FT FT US-09-889-314-2 (1-496) x AAV03553 (1-15620)
FT FT
FT FT 17 AsnIleMetSerGlnValLeuThrSerThrProGlnGlyValProGlnGln----- 33
FT FT Db 9936 AATCTGTGGCAGATCGACGCTACGCTACGCCCATTCATGATGCCCAAGGACGATTA 9995
FT FT
FT FT OY 34 ---AspLysIleSerGlyAsn-GluThr-----LysGlnIleGlnGln 47
FT FT Db 9996 ACATCGAAGCCGGCGGCGGAAATTTGACACTCTATGCCGTAGAGAGCTCACTACGACAAAC 10055
FT FT OY 47 FArgGlnGlyLysAsnThrGlnMetGlnUserAspAlaThr-IleAlaGlyAlaSerGlyL 67
FT FT

```

0y	483	eserclYAlaIlealagly-AtaIstIsYFthAsnSn	495
	111	***** 111111***** 111	
Db	1648	GAGTGGCTGACGCTCAGCAAGCACATCAGCTTCTGAT	1685
RESULT 16			
ID	AAV52497	standard; DNA; 2550 BP.	
XX	AAV52497		
AC	AAV52497;		
XX			
DT	23-OCT-1998	(first entry)	
XX			
DE	Streptococcus pneumoniae genome fragment SEQ ID NO:364.		
XX			
KW	Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;		
KW	computer readable medium; vaccine; pharmaceutical composition; ds.		
XX			
OS	Streptococcus pneumoniae.		
PN	MO9818931-A2.		
PN			
PD	07-MAY-1998.		
XX			
PF	30-OCT-1997;	97WO-US19588.	
XX			
PR	31-OCT-1996;	96US-0029960.	
XX			
PA	(HUMA-) HUMAN GENOME SCI INC.		
PI	Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;		
PI	Kunsch CA, Rosen CA;		
XX			
DR	WPI: 1998-272225/24.		
XX			
PT	Computer-readable medium with recorded Streptococcus pneumoniae		
PT	polynucleotide sequences - useful in diagnostic kits and assays, and		
PT	pharmaceutical compositions and vaccines for Streptococcus		
PT	pneumoniae		
XX			
XX			
PS	Claim 1; Page 1379-1380; 1409pp; English.		
XX			
CC	The present invention describes a computer readable medium which has		
CC	the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52524)		
CC	recorded on it, or a representative fragment or a sequence at least 95%		
CC	identical to SEQ ID NO:1 to 391. The nucleotide sequences depicted in		
CC	SEQ ID NO:1 to 391 (AAV52134 to AAV52524) are genomic fragments from		
CC	Streptococcus pneumoniae. The present invention also describes an		
CC	isolated nucleic acid molecule encoding a homologue of any of the		
CC	fragments of the S.pneumoniae genome (SEQ ID NO:1 to 391) where the		
CC	nucleic acid molecule is produced by a process comprising: (a) screening		
CC	a genomic DNA library using as a probe a target sequence defined by any		
CC	of the sequences in SEQ ID NO:1 to 391, identifying members of the		
CC	library which contain sequences that hybridise to the target sequence and		
CC	isolating the nucleic acid molecules from the members; or (b) isolating		
CC	MRNA, DNA or cDNA produced from an organism, amplifying nucleic acid		
CC	molecules whose nucleotide sequence is homologous to amplification		
CC	primers derived from the fragment of the S. pneumoniae genome to prime		
CC	the amplification and isolating the amplified sequences. The computer		
CC	readable medium can be used in a computer-based system for identifying		
CC	fragments of the S. pneumoniae genome of commercial importance, or		
CC	expression modulating fragments of the S. pneumoniae genome. Products		
CC	from the present invention can be used in diagnosis kits and assays, and		
CC	pharmaceutical compositions and vaccines for S. pneumoniae.		
XX			
SO	Sequence 2550 BP; 630 A; 837 C; 584 G; 498 T; 1 other;		
Alignment Scores:			
Pred. No.:	4,69e-05	Length:	2550
Score:	192.00	Matches:	110
Percent Similarity:	33.97%	Conservative:	68
Best Local Similarity:	20.99%	Mismatches:	275
Query Match:	8.10%	Indels:	71

[illegible]

D	557	TCAGCTCCGCGTCACACAGCGCCCTCGGCGCTCAGCAAGATATCTCAGCGCTTGAAATCGCA	616
O	127	SerLysSerMetGluSerThrLeuGlnSerLeuGlnSerLeuSerAlaIleGlnMetLys	146
D	617	TCAAACAGAGCGCTCGCGCTTCAGACATCAACAGGATGCATCACTCCAGCAAGCAACGATCGC	676
O	147	GluValaGluAlaValValValAlaIleLeuSerGlyLysSerSerGlySerAlaLysLeu	166
D	677	TCGGCCTCAGCAAGCAACAGCGCCGCTCTGAATCCGATCAACCAAGTGGCTCAGGCTCAGCA	736
O	167	GluThrProGluLeuProLysProGlyValThrProArgSerGluValIleGluIleGly	186
D	737	AGTACCTCAGCATTCGAATCAGCATCAACAGTGCATCGCGCTTCAGCAAGCA-----	790
O	187	LeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGluAlaThrLysSerAlaLeuSerAsn	206
D	791	AGTGGTTCAGCCCTCAGCAACCAATCTCAGCGCTTGAAATCGGCAATCAACAGATCGCTCCGT	850
O	207	TyrAlaSerThrGlnAlaGlnAlaIleAspGlnThrAsnLys-LeuGlyLeuGlnGlnAla	226
D	851	TCAGCAAGTAACTAGCGCTCAGCATCAGCGTCAACAAAGTCTCGGCTTCACAGCGTCAAGC	910
O	226	AlaLeuLysLeuAspLysGlu-----ArgGluGluTyrGlnGluMetLysAlaAlaGlu	243
D	911	AGTGGCTTGAGTCATCATCAACGAGTACGTGACCTCAGCGTCAAGC-ACATCAGCTTCTGA	969
O	243	GluThrLysSerLysAspLeuGlnGluGlyThrMetAspThrValAsnThrValMetIleAla	263
D	970	ATCTGCATCAACCAAGTGGCGTCACTCAGCATCAACAGCGCTCAGCTTCAGCAAGTAC	1029
O	263	LserValaAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyValaGlyLe	283
D	1030	CAATGGCTCA-----GCTTCAGCAAGTACCAAGTCTTCACCGTC	1068
O	283	ValaGlyLeuAlaAlaGlyAlaAlaValaGlyAlaAlaAlaAlaGlyGlyAlaAlaGlyAla	303
D	1069	AGCGTCGACAAAGTGGCGGCTTCACCAAGTGCATGATCGGATCGGATCAACAGTGCCTGC	1128
O	303	AlaAlaAlaThrThrValAlaIleThrGlnIleThrValGlnAlaValaGlnAlaValaLys	323
D	1129	AGCGTCGACAAAGTCTGAGCGCC-----TCAGCTCAGCAATCAACAGTGCCTGC	1179
O	323	SglnAlaValIlePheThrAlaValaArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLys	343
D	1180	TTTCAGCAAGTACTGTGCATCAGCATCAGCATCAACGATGACA--TCGCGCTTCAGCAAG	1236
O	343	SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys	363
D	1237	TACAGCGCGCTCAGCTTCAGCAACCAAGTGCCTCAGCGCTCAAGTACGATCAGCGCTC	1296
O	363	SgLYleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLe	383
D	1297	AGCTCAGCAAGCAACAGTGGCTCAGGCTCA-----GCAAGTACCAAGTGGCTCAGC	1347
O	383	uSerLysValIleSerSerLeuThrSerLysTrpValThrValaGlyValaVala	403
D	1348	CTCAGCGCTGCAAGATGGCGTGGCTTCAGCAAGTACCTCAGCGCTCGAATCAGCATCAAC	1407
O	403	LAlaIleProAlaLeuGlnGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGlnAsn	423
D	1408	CAGTGCATCAGCTTCAGCATCAACAAAGTCTTCAGCTTCACCAAGTACCAAGTACGCTGGCG	1467
O	423	nValaIaGlnPheGlnLysGluValaGlyLysLeuGlnAlaIleAlaIleAspMetIleSerMe	443
D	1468	TTTCAGCATCAACGAGTGGCTTCAGCTCAGCGTCAACAGTGCCTTCGATCAGCATCAAC	1527
O	443	LpheThrGlnPheTrpGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyGlnSerAs	463
D	1528	AAGTGCCTCGGCTTCAGCAACCAAGTGGCTTCAGCGCTTCACCAAGTACTAGTGCATCGCG	1587
O	463	nGluMetThrGlnLysAlaThrLysLeuGlyValaGlnIleLeuLysAlaTyrAlaAlaIle	483
D	1588	TTTCAGCATCAACGAGTGGCTTCAGTCCGATCAACGAGTGGCTTCGCTTCAGCATCAAC	1647

CC library which contain sequences that hybridise to the target sequence and
 CC isolating the nucleic acid molecules from the members; or (b) isolating
 CC mRNA, DNA or cDNA produced from an organism, amplifying nucleic acid
 CC molecules whose nucleotide sequence is homologous to amplification
 CC primers derived from the fragment of the S. pneumoniae genome to prime
 CC the amplification and isolating the amplified sequences. The computer
 CC readable medium can be used in a computer-based system for identifying
 CC fragments of the S. pneumoniae genome of commercial importance, or
 CC expression modulating fragments of the S. pneumoniae genome. Products
 CC from the present invention can be used in diagnosis kits and assays, and
 CC pharmaceutical compositions and vaccines for S. pneumoniae.

XX Sequence 4483 BP; 1092 A; 1451 C; 1051 G; 884 T; 5 other;

Alignment Scores:

Pred. No.:	2.76e-05	Length:	4483
Score:	199.00	Matches:	111
Percent Similarity:	37.80%	Conservative:	78
Best Local Similarity:	22.20%	Mismatches:	274
Query Match:	8.40%	Indels:	39
DB:	19	Gaps:	11

US-09-889-314-2 (1-496) x AAV52496 (1-4483)

QY 2 ThrAsmMetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGln 21
 DB 689 ACAAGTCGCTCAGCCTCAGCAAGTACTGATGCATCAGCTTCAGCATACAGAGTCATCG 748
 QY 22 ValLeuThrSerThrProGlnGlyValProGlnAspLysLeuSerGlyAsnGluThr 41
 DB 749 GCTTCGGCGTCACCC-----AGTCGATCAGAGTCA 778
 QY 42 LysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGluSerAspAlaThrIle 61
 DB 779 GCAAGTACCAAGTCGCTCAGCTCCGATCAACAGTCCTCGGCTTCAGCAAGCAGCT 838
 QY 62 AlaGlyLysSerGlyLysAspLysThrSerSerThrLysThrGluThrAlaProGln 81
 DB 839 GCGTCGCTTCAGCAAGTACTAGCGCCTCAGCCTCAGCCTCAGCAGTGCCTCAGCA 898
 QY 82 GlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGly 101
 DB 899 GCAAGTATTCAGCGCTCGATCGCATCAACAGAGTCGCTCGCTTCAGCAAGTACTAGC 958
 QY 102 ValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThr 121
 DB 959 GCTTCAGCCTCAGCCTCAGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 1012
 QY 122 SerIleGluGluAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeuSer 141
 DB 1013 -----GATCGCATCAACAGTGCCTCGCTTCAGCAAGTACTAGCGCTCAGCCTCA 1066
 QY 142 AlaAlaGlnMetLysGluValAlaGluAlaValAlaAlaAlaLeuSerGlyLysSerSer 161
 DB 1067 GCGTCAACAAAGTATCGCTTCAGCATCAACAGAGTGCCTCGCTTCAGCAAGTACTAGC 1126
 QY 162 GlySerAlaLysLeuGluThrProGluLeuProLysProGlyValThrProAlaSerGlu 181
 DB 1127 GCTTCAGCCTCAGCGTCAACA---AGTCATCGCTTCAGCGTCAACAGTGCCTCAG 1183
 QY 182 ValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaAlaThrLys 201
 DB 1184 TCAGCATCAACAGT---GCGTCAGCCTCAGCAACAGCATACATCAGCTTCGATTCGATCA 1240
 QY 202 SerAlaLeuSerAsnTyraLysSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGly 221
 DB 1241 ACCAGTGCCTCAGCCTCAGCATTCAGCAGCGCTCAGCTTCAGCAAGTCAAGTGCCTCA 1300
 QY 222 LeuGluLysGlnAlaIleLysIleAspLysGlu-----ArgGluGluIleGlnGln 238
 DB 1301 GCTCAGCGCTCAGCAAGTCTGCTTCAGCAAGTCAAGTCAAGTCAAGTCAAGTCAAGT 1359
 QY 239 MetLysAlaIleGluGlnLysSerLysAspLeuGluGlyThrMetAspThrValAsnThr 258

DB 1360 ACCAGTGCCTCAGCCTCAGCATCAGAGTGCCTCAGCAACAGTGCATCTGAATCG 1419
 QY 259 ValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIleThr 278
 DB 1420 GCATCAACAGTGCCTCAGCCTCAGCATCAGTACT---AGGCGCTCAGCCTCAGCATCAAG 1476
 QY 279 CysGlyValGlyLeuValLeuAlaIleGluAlaIleValGlyAlaAlaAlaIleGly 298
 DB 1477 AGTGCCTCAGCCTCAGCAGTACTGCTCAGCTTCAGCTTCAGCAAGTCAAGTCAAGTCAAG 1536
 QY 299 GlyAlaAlaGlyAlaAlaAlaAlaAlaThrThrValAlaIleThrGlnIleThrValGlnAlaVal 318
 DB 1537 TCAGCGTCAACAGCGCCTCAGCT---TCAGCAAGTCAACAGTGCCTCAGCCTCAGCGTGC 1593
 QY 319 ValGlnAlaValLysGlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAlaIle 338
 DB 1594 ACAAGTGCCTCAGCCTCAGCATCAGTACTCAGCTTCGATCAACATCAACATCAAGTGC--- 1650
 QY 339 LysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAla 358
 DB 1651 TCGGCTTCAGCATCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1710
 QY 359 Lys---AlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAla 377
 DB 1711 AGTACCAAGTGCCTCAGCCTCAGCATCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCC 1770
 QY 378 LysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysThrValAlaThrVal 397
 DB 1771 TCTGAATCCCATCAACATCAAGTGCCTCAGTCAAGCAAGCAAGTGCCTCAGTCAAGC 1830
 QY 398 GlyValGlyValValAlaAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeu 417
 DB 1831 TCAACAGTGCCTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1890
 QY 418 SerGlnMetGlnGlnAsnValAlaGlnPheGlnLysGlnValGlyLys-LeuGlnAlaAl 437
 DB 1891 CTGAATCTGCATCAACAGTGCCTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1944
 QY 437 AlaAspMetIleSerMetThrGlnPheThrGlnIleAlaSerLysIleAlaSerLys 457
 DB 1945 AGCA-----AGTACCAAGTGCCTCAGCCTCAGC 1971
 QY 457 sGlnThrGlyLysSerAsnGlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLe 477
 DB 1972 ATCAACCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 2031
 QY 477 uLysAlaTyraLysAlaIleSerGlyAlaIleAlaGlyAlaHisLysThrAsnAsnPe 496
 DB 2032 CTCGCTTCAGCAAGTCAAGTGCCTCAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 2088
 RESULT 14
 AAA05548
 ID AAA05548 standard; DNA: 3633 BP.
 XX AAA05548;
 AC
 XX 24-MAY-2000 (first entry)
 DE Streptococcus pneumoniae type 4 nucleotide sequence 4181.2.
 XX Streptococcus pneumoniae; vaccine; screening; protein antigen;
 KW antibacterial; antiinflammatory; meningitis; infection; diagnosis;
 KW pneumococcal disease; ds.
 XX Streptococcus pneumoniae.
 OS
 XX MO200006737-A2.
 PN
 XX 10-FEB-2000.
 PD
 XX 27-JUL-1999; 99WO-GB02451.
 PF
 XX

```

Db 4668 GCTTCAGCAAGTACC-----TCAGCGCTCGAATCA 4717
Oy 42 LysGlnIleGlnInThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThrIle 61
Db 4718 GCATCAACAACTGGCTGGCTTCAGCATCAACGAGTCATCAGCTTCAGCATCAACAACT 4777
Oy 62 AlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGln 81
Db 4778 GCTTCAGCTTCAGCAAGTACAGTCGCGCTTCAGCATCAACGAGTCGCTTCAGTCA 4837
Oy 82 GlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGly 101
Db 4838 GCCTCAACAGGCTCGCTCGAATCCGATCAACAGTCGCTTCAGCAAGCACAGT 4897
Oy 102 ValSerGlyAlaAlaAlaThr-----ThrAlaSerAsnThrAlaThrIle 117
Db 4898 GCTTCGCTTCAGCATCAACGAGTCGCTTCAGTCATCAACGAGTCGCTTCAGCTCA 4957
Oy 118 AlaMetGlnThrSerIleGluGlnAlaSerLysSerMetGluSerThrLysGluSerLeu 137
Db 4958 GCACACATCATCAGCTTCGAATCTGCATCAACCACTGCGTCACCTCCGACATCAACAAGC 5017
Oy 138 GlnSerLeuSerAlaAlaGlnMetLysGlnValGlnAlaValAlaAlaAlaLeuSer 157
Db 5018 GCCTCGGCTCAGCAAGTACAAAGTCTTCAGCCTCAGCATCAACCAAGTCATCAGCTTCA 5077
Oy 158 GlyLysSerSerGlySerAlaLysLeuGlnThrProGlnLeuProLysProLysValThr 177
Db 5078 GCCTCAACAGGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAGCTTCAG 5137
Oy 178 ProArgSerGluValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGly 197
Db 5138 GCGTCAGCTTCAGCAACGACCAAGTCGTCAGCTTCAGCATCAACGAGTCCTCGGCTTCG 5197
Oy 198 GlnAlaThrLysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThr 217
Db 5198 GCATCAACAGGCTTCAGCATCA-----GCATCAACAGGCTTCAGCTTCAGCAAGT 5251
Oy 218 AsnLysLeuGlyLeuGlnLysGlnAlaIleLysIleAspLysGlnArgGlnLysGln 237
Db 5252 ACT-----AGTCATCAGCATCAGCATCAACCAAGTCATCAGCTTCAGCA 5296
Oy 238 GlnMetLysAlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsn 257
Db 5297 AGTATCTCAGCTTCGAATCGGCATCAACGAGTCATCAGCATCAACCAAGTCAGTGA 5356
Oy 258 ThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePhe 277
Db 5357 TCGGCTTCAGCTCAACGAGTCATCAGTCATCAGCAACGACCAAGTCGCTTCAGCA 5416
Oy 278 ThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAla 297
Db 5417 TCACGAGTGGCTTCAGCTTCAGCAAGTATCTCAGCTCGAATCGGCATCAACGAGTGG 5476
Oy 298 GlyGlyAlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 317
Db 5477 TCAGGCTTCAGCAAGTACGATCGCTTCAGCAACGACCAAGTCGCTTCAGCA 5536
Oy 318 ValValGlnAlaValLysGlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAla 337
Db 5537 TCACACAGTGGCTTCAGCTTCAGCAAGTATCTCAGCTCGAATCGGCATCAACGAGTGG 5596
Oy 338 IleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIle 357
Db 5597 ---TCAGCTTCAGCAAGTACGATCGATCAGTCATCAGCAACGAGTCATCGGCTTCA 5653
Oy 358 AlaLysAlaIleSerLysGlyLysSerLysValPheAlaLysGlyThrGlnMetIleAla 377
Db 5654 GCAAGTACAGGCGCTTCAGCTTCAGCAACGACCAAGTCGCTTCAGC----- 5698
Oy 378 LysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysTrpValThrVal 397
Db 5699 -----TCAGCAAGTACAGGCTTCAGCTTCAGCAACGACCAAGTCAGTGC 5740

```

```

Oy 398 GlyValGlyValValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeu 417
Db 5741 TCAGCTTCAGCAAGTACAGTCGCTTCAGCTTCAGCATCAACGAGTCCTTCAGCTTCG--- 5797
Oy 418 SerGlnMetGlnInAsnValAlaGlnPheGln---LysGlnValGlyLysLeuGlnAla 436
Db 5798 -----CCTCAACAGTGGCTTCAGCTTCAGCTTCAGCAACGAGTCGCTTCAGCA 5848
Oy 437 Ala-----AlaAspMetIleSerMetPheThrGlnPheTrpGln 449
Db 5849 GCACAGTGGCTTCAGCTTCAGCAACGACCAAGTCGCTTCAGCTTCAGCTTCAGCTTCAG 5902
Oy 450 GlnAlaSerLysIleAlaSer-----LysGln 458
Db 5903 GT-GGCTCAGCTTCAGCAAGTACATCAGTTCAATTCAGCAACCAATTCGAATCAGCA 5961
Oy 459 ThrGlyLysSerAsnGlnMetThrGlnLysAlaThrLys----- 471
Db 5962 GTTGAATACTTCTGATTCGACAGTAAATCCAAAAAGAAATTCCTAATACAGTACT 6021
Oy 472 -----LysGlyAlaGlnIleLeuLysAlaThrAlaAlaIleSerGly 485
Db 6022 GAGTCGTCATTTGATCTGTACTTGGAGTTCAGCAAGCTTCAGCAAGT 6072

RESULT 13
AAV52496
ID AAV52496 standard; DNA; 4483 BP.
XX
AC AAV52496;
XX
DT 23-OCT-1998 (first entry)
XX
DE Streptococcus pneumoniae genome fragment SEQ ID NO:363.
XX
KW Streptococcus pneumoniae; S. pneumoniae; genome; diagnosis; assay;
KM computer readable medium; vaccine; pharmaceutical composition; ds.
XX
OS Streptococcus pneumoniae.
XX
PN W09818931-A2.
XX
PD 07-MAY-1998.
XX
PF 30-OCT-1997; 97WO-US19588.
XX
PR 31-OCT-1996; 96US-0029960.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Barash SC, Choi GH, Dillon PJ, Dougherty BA, Fannon M;
XX
PI Kunsch CA, Rosen CA.
XX
DR WPI: 1998-272225/24.
XX
PT Computer-readable medium with recorded Streptococcus pneumoniae
PT polynucleotide sequences - useful in diagnostic kits and assays, and
PT pharmaceutical compositions and vaccines for Streptococcus
PT pneumoniae
PS
PS Claim 1: Page 1376-1378; 1409pp; English.
XX
XX The present invention describes a computer readable medium which has
XX the nucleotide sequences SEQ ID NO:1 to 391 (AAV52134 to AAV52574)
XX recorded on it, or a representative fragment of a sequence at least 95%
XX identical to SEQ ID NO: 1 to 391. The nucleotide sequences depicted in
XX SEQ ID NO:1 to 391 (AAV52134 to AAV52574) are genomic fragments from
XX Streptococcus pneumoniae. The present invention also describes an
XX isolated nucleic acid molecule encoding a homologue of any of the
XX fragments of the S. pneumoniae genome (SEQ ID NO:1 to 391) where the
XX nucleic acid molecule is produced by a process comprising: (a) screening
XX a genomic DNA library using as a probe a target sequence defined by any
XX of the sequences in SEQ ID NO:1 to 391, identifying members of the

```


CC with the expression vector pAD431 to give pCPN533alpha. The
CC plasmid was used to transform an E. coli host, which was cultured
CC to give the antigenic polypeptide, polypeptide A. Polypeptide A
CC and primers and probes derived from its DNA can be used in assays
CC for the detection of polypeptide A antibodies and DNA,
CC respectively, useful in the diagnosis of C. pneumoniae infection.

XX Sequence 813 BP; 286 A; 181 C; 187 G; 159 T; 0 other:

Alignment Scores:

Pred. No.:	5,58e-84	Length:	813
Score:	1245.00	Matches:	259
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	52.55%	Indels:	0
DB:	17	Gaps:	0

US-09-889-314-2 (1-496) x AAT14613 (1-813)

```
QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23
Db 1 ATGCTATTTCATCTCTTTCAGGACCTGACATCAAAAAAATATCATGTCTCAAGTCTG 60
QY 24 ThrSerThrProGlnGlnValProGlnLysLysLeuSerGlyAsnGlnThrLysGln 43
Db 61 ACATCGACACCCGAGGCGTCCCAACACAGATAGCTGTGTGCGAACGAAACGACGAA 120
QY 44 IleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63
Db 121 ATACGCAACACGTCAGGCTAAACACTGAGATGGAACGATGCGCATATTGCTGCT 180
QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83
Db 181 GCTCTGGAAAAAGACAAACTCTCGACTACAAAACGAAACGACGCTCCACACAGGGA 240
QY 84 ValAlaIleGlyLysGlnSerSerSerGlnLysAlaGlyAlaAspThrGlyValSer 103
Db 241 GTGTGCTGTGGGAAGAAATCTCTCAAAAGTCAAAAGCAGGTGCGATCGAGTATCA 300
QY 104 GlyAlaIleAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
Db 301 GGAGCGGCTGCTACTACACGATCAAAATCTGCAACAAAATTTGCTATGACGCTCTATT 360
QY 124 GlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaIle 143
Db 361 GAGAGGCGAGCAAAAGATGAGTCTTACCTTAGAGTCACTTCAAGCCTCAGTCCGCG 420
QY 144 GlnMetLysGlnValGlnAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlySer 163
Db 421 CAATGAAAGAGTGAAGCGGTGTGTGCTGCCCTCTCAGGGAAAGTTCCGGGTTCC 480
QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlnValThrProArgSerGlnValIle 183
Db 481 GCAAAATTTGGAAACCTGAGTCCCAAGCCGGGGTGACACCAAGATCAAGAGTTATTC 540
QY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThrLysSerAla 203
Db 541 GAAATCGGACTCGCGCTTCTTAAACCAATTCAGACATTGGGAAAGCCCAAAATCTGCC 600
QY 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnLeu 223
Db 601 TTATCTAACTATGCAAGTACACACAGCACAAAGCACAACAAATTAAGTCTAGAA 660
QY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnIleMetLysAlaIleAlaGln 243
Db 661 AAGCAAGCATATAAAATCATTAAGAAAGCAAGAAATATCAAGAGTGAAGCTGCCGAA 720
QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAla 262
Db 721 CAGAACTCTAAGATCTGCAAGGAACAATGATGATCTGATGATCGCG 777
```

RESULT 10
AAT14619

ID AAT14619 standard; DNA; 1296 BP.

XX AAT14619;

DT 11-NOV-1996 (first entry)

DE DHFR/C. pneumoniae antigen variant fusion protein DNA.

KW Polypeptide antigen; strain YK41; plasmid; probe;

KW pCPN533T; primer; assay; detection; antibody; diagnosis;

KW infection; fusion protein; dihydrofolate reductase; DHFR;

OS variant; Chlamydia pneumoniae; ds.

XX Synthetic.

FT Key Location/Qualifiers

FT CDS 1..1296

FT /tag= a

XX /note= "STOP codon absent"

XX W09609320-A1.

XX 28-MAR-1996.

XX 20-SEP-1995; 95WO-JP01896.

XX 28-APR-1995; 95JP-0106011.

XX 20-SEP-1994; 94JP-0224711.

XX 28-APR-1995; 95JP-0106006.

XX 28-APR-1995; 95JP-0106008.

XX 28-APR-1995; 95JP-0106009.

XX 28-APR-1995; 95JP-0106010.

XX (HTTB) HITACHI CHEM CO LTD.

XX Iutsu H, Matsumoto A, Obara K;

XX WPI: 1996-188399/19.

XX P-PSDB; AAR94585.

XX Claim 26; Pages 103-107; 128pp; Japanese.

XX The present sequence encodes the dihydrofolate reductase (DHFR)/

XX C. pneumoniae polypeptide antigen variant fusion protein.

XX C. pneumoniae strain YK41 was cultured and genomic DNA extracted

XX to prep. a lambda gt11 DNA library. The library was then screened

XX with an anti-YK41 monoclonal antibody (Mab), which was prepd. by

XX fusing spleen cells from a mouse infected with YK41 with myeloma

XX p3/NSI/1-Ag4-1 to produce a Mab expressing hybridoma. The DNA

XX obtd. was then fused with DHFR DNA and the expression vector

XX pAD431 to give pCPN533T. The plasmid was used to transform an

XX E. coli host, which was cultured to give an antigenic polypeptide

XX fusion protein. The fusion protein and primers and probes derived

XX from its DNA can be used in assays for the detection of the

XX antigenic polypeptide antibodies and DNA, respectively, useful in

XX the diagnosis of C. pneumoniae infection.

XX Sequence 1296 BP; 407 A; 300 C; 324 G; 265 T; 0 other:

XX Alignment Scores:

Pred. No.:	9.68e-84	Length:	1296
Score:	1245.00	Matches:	259
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	52.55%	Indels:	0
DB:	17	Gaps:	0

US-09-889-314-2 (1-496) x AAT14619 (1-1296)

```

XX  JP08304404-A.
XX
XX  22-NOV-1996.
XX
XX  28-APR-1995; 95JP-0106014.
XX
XX  28-APR-1995; 95JP-0106014.
XX
XX  (HITB ) HITACHI CHEM CO LTD.
XX
XX  WPI; 1997-056178/06.
XX  P-PSDB; AAM01743.
XX
XX  Detection and determination of anti-Chlamydia pneumoniae antibody -
XX  using the polypeptide C as the antigen
XX
XX  Example 3; Page 17-18; 18pp; Japanese.
XX
XX  The gene encodes the Chlamydia pneumoniae 53 kDa antigen (AAM01743). A
XX  method for the detection and determination of anti-C. pneumoniae
XX  antibodies in a sample comprises using at least 5 consecutive amino
XX  acids of the polypeptide C 73 kDa antigen (AAM01742).
XX
XX  Sequence 1048 BP; 360 A; 241 C; 231 G; 216 T; 0 other;
SQ

```

Alignment Scores:

Pred. No.:	4,09e-85	Length:	1048
Score:	1262.00	Matches:	262
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	53.27%	Indels:	0
DB:	18	Gaps:	0

US-09-889-314-2 (1-496) x AAT59311 (1-1048)

```

OY  1 AspThrAsnMetSerIleSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20
    |||||
DB  227 GATCAACAACATGCTATTTCATCTTCTTCAGACCTGACACATATAAATATCATGTCT 286
    |||||
OY  21 GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGln 40
    |||||
DB  287 CAAGTCTGACATCGACACCCCGAGCGCTGCCCAACAAGATTAAGCTGCTGCAACGAA 346
    |||||
OY  41 ThrGlyGlnIleGlnIleThrArgGlnGlyLysAsnThrGlnMetGluSerAspAlaThr 60
    |||||
DB  347 ACCGAACCAATATACAGAAACACGCTCAAGGTAAACACAGATGGAAGACCGATGCCACT 406
    |||||
OY  61 IleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaPro 80
    |||||
DB  407 ATTGCTGCTGCTTCTGGAAGAAAGCAAAACTCTCTCCACTACAAAACACAGAACGCTCCA 466
    |||||
OY  81 GlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThr 100
    |||||
DB  467 CAACAGGAGAGTTCCTGCTGGAAAGATCCTCAGAAAGTCAAAAGCAGAGTCTGATAC 526
    |||||
OY  101 GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120
    |||||
DB  527 GGAGTATCAGAGAGCGGCTCTACTACAGCATCAAAATACCAAAATGCTATGCA 586
    |||||
OY  121 ThrSerIleGlnGlnLysSerLysSerMetGluSerThrLeuGlnSerLeuGlnSerLeu 140
    |||||
DB  587 ACCCTGATTGAAAGCGCAGCAAAAGTATGAGTCTTACCTTAGAGTCACCTTCAAGCCTC 646
    |||||
OY  141 SerAlaIleGlnMetLysGluValGluAlaValAlaAlaAlaLeuSerGlyLysSer 160
    |||||
DB  647 AGTGGCGCCCAATGAAAGCGAAGCGGTGTGTGCTGCCCTCAGAGGAAAG 706
    |||||
OY  161 SerGlySerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSer 180
    |||||
DB  707 TCGGGTTCCGCAAAATGGAACACCTGAGCTCCCAAGCGCGGGTGAACCAACATCA 766
    |||||
OY  181 GluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnIleThrLeuGlnGluAlaThr 200
    |||||

```

```

DB  767 GAGGTATCGAATCGACGCTCGCTTCTTAAGCAATTCAGACATTTGGAGAGCCACA 826
    |||||
OY  201 LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
    |||||
DB  827 AAATCTGCCTTATCTATCATCTATGCAAGTACCAAGCACAAGCAGACCAATCAATAACTA 886
    |||||
OY  221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGluArgGluGlnGlyTyrGlnGlnMetLys 240
    |||||
DB  887 GGTCTAGAAAGCGACGATTAATAATCGATTAAGAGAGAGAAATACCAAGAGATGAAG 946
    |||||
OY  241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMet 260
    |||||
DB  947 GCTGCCGACAGACAGCTTAAGATCTCGAAGGACAAATGATATCTCAATCTGTGANG 1006
    |||||
OY  261 IleAla 262
    |||||
DB  1007 ATCGCG 1012
    |||||

```

RESULT 9

AAT14613

ID AAT14613 standard; DNA; 813 BP.

XX AAT14613;

AC AAT14613;

XX 07-NOV-1996 (first entry)

DE C. pneumoniae polypeptide antigen (polypeptide A) DNA variant.

XX

XX Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;

KW PCP533alpha; primer; assay; detection; antibody; diagnosis;

KW Infection; variant; ds.

XX

XX Chlamydia pneumoniae.

OS

XX

XX Key Location/Qualifiers

FT 1..812

FT CDS /*tag= a

FT /*note= "STOP codon absent"

XX

XX W09609320-A1.

XX

XX 28-MAR-1996.

XX

XX 20-SEP-1995; 95WO-JP01896.

XX

XX 28-APR-1995; 95JP-0106011.

XX

XX 20-SEP-1994; 94JP-0224711.

PR 20-SEP-1994; 94JP-0224711.

PR 28-APR-1995; 95JP-0106006.

PR 28-APR-1995; 95JP-0106008.

PR 28-APR-1995; 95JP-0106009.

PR 28-APR-1995; 95JP-0106010.

PR 28-APR-1995; 95JP-0106010.

XX

XX (HITB) HITACHI CHEM CO LTD.

XX

PI Izutsu H, Matsumoto A, Obara K;

XX

XX WPI; 1996-188399/19.

DR P-PSDB; AAR54580.

XX

XX Recombinant Chlamydia pneumoniae antigen and antibodies to it

PT used for detection and assay of C. pneumoniae e.g. in clinical

PT diagnosis

XX

XX Claim 10; Pages 64-66; 128pp; Japanese.

XX

XX The present sequence encodes a variant of the C. pneumoniae

CC polypeptide antigen, polypeptide A. C. pneumoniae strain YK41 was

CC cultured and genomic DNA extracted to prep. a lambda gt11 DNA

CC library. The library was then screened with an anti-YK41

CC monoclonal antibody (Mab), which was prep. by fusing spleen cells

CC from a mouse infected with YK41 with myeloma P3/NSI/1-Ag4-1 to

CC produce a Mab expressing hydridoma. The DNA obtd. was then fused

AA14622
ID AA14622 standard; DNA: 1048 BP.
AC AA14622;
XX
DT 11-NOV-1996 (first entry)
DE C. pneumoniae polypeptide antigen (polypeptide A) DNA clone 53-35.
XX
KM Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;
KW pcPN533alpha; primer; assay; detection; antibody; diagnosis;
XX Infection; clone; ds.
OS Chlamydia pneumoniae.
XX
FH Key Location/Qualifiers
FT 236..1012
FT CDS /*tag= a
FT /note= "STOP codon absent"
XX
PN WC0609320-A1.
XX
PD 28-MAR-1996.
XX
PE 20-SEP-1995; 95WO-JP01896.
XX
PR 28-APR-1995; 95JP-0106011.
PR 20-SEP-1994; 94JP-0224711.
PR 28-APR-1995; 95JP-0106006.
PR 28-APR-1995; 95JP-0106008.
PR 28-APR-1995; 95JP-0106009.
PR 28-APR-1995; 95JP-0106010.
XX
PA (HITB) HITACHI CHEM CO LTD.
XX
PI Izutsu H, Matsumoto A, Obara K;
XX
DR WPI; 1996-188399/19.
DR P-PSDB; AAR94586.
XX
PT Recombinant Chlamydia pneumoniae antigen and antibodies to it
PT used for detection and assay of C. pneumoniae e.g. in clinical
PT diagnosis
XX
PS Example 1; Pages 75-77; 128pp; Japanese.
XX
XX The present sequence encodes the C. pneumoniae polypeptide antigen
CC polypeptide A clone, 53-35. C. pneumoniae strain YK41 was
CC cultured and genomic DNA extracted to prep. a lambda gill DNA
CC library. The library was then screened with an anti-YK41
CC monoclonal antibody (mAb), which was prepd. by fusing spleen cells
CC from a mouse infected with YK41 with myeloma P3/NS1/J-Ag4-1 to
CC produce a MAb expressing hybridoma. The DNA obcd. was then fused
CC with the expression vector pAD431 to give pcPN533alpha. The
CC plasmid was used to transform an E. coli host, which was cultured
CC to give the antigenic polypeptide, polypeptide A. Polypeptide A
CC and primers and probes derived from its DNA can be used in assays
CC for the detection of polypeptide A antibodies and DNA.
CC respectively, useful in the diagnosis of C. pneumoniae infection.
XX
SQ Sequence 1048 BP; 360 A; 241 C; 231 G; 216 T; 0 other;

Alignment Scores:
Pred. No.: 4 09e-85 Length: 1048
Score: 1262.00 Matches: 262
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 53.27% Indels: 0
DB: 17 Gaps: 0

US-09-889-314-2 (1-496) x AA14622 (1-1048)

OY 1 AspThrAsnMetSerIleSerSerSerGlyProAspAsnGlnIlyAsnIleMetSer 20

Db 227 GATACAAACATGCTATTTCATCTTCTTACAGACCTCGAACATCAAAAAATATCATGCTCT 286
OY 21 GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlu 40
Db 287 CAAGTCTCTGACATCGACACCCAGGGCGTCCGCCAACAAGATAGCTGCTGGCAACGAA 346
OY 41 ThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThr 60
Db 347 ACCAAGCAAAATACAGCAAAACACGTCAGGTAAACACGAGATGGAAGCCATGCCACT 406
OY 61 IleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaPro 80
Db 407 ATTGCTGGTCTCTTCTGAAAAGCAAACTCTCTCGATACAAAACAGAAACAGCTCCA 466
OY 81 GlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThr 100
Db 467 CAACAGGAGACTGCTCTGCGGAAAGATCTCGAAGAAAGTCAAAAGGACAGTCTGATACT 526
OY 101 GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120
Db 527 GGAGTATCAGAGCGGCTGCTACTACAGCATCAAAATACGCAAAATAATGCTATGACG 586
OY 121 ThrSerIleGlnGlnAlaSerLysSerMetGluSerThrLeuGlnSerLeuSerLeu 140
Db 587 ACCCTATTGAAAGAGGAGCAAAAGATAGTACCTTACCTTACCTTCAAAAGCCTC 646
OY 141 SerAlaAlaGlnMetLysGluValGluAlaValAlaValAlaAlaLeuSerGlyLysSer 160
Db 647 AGTCCCGCCGCAAAATGAAAGAGCGAGCGGTGTGTGCTCTCCCTCTCAGGGAAGT 706
OY 161 SerGlySerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSer 180
Db 707 TCGGGTTCGCAAAATGGAACACCTGACTCCCAAGCCGCGGGTGACCAAGATCA 766
OY 181 GlnValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGluAlaThr 200
Db 767 GAGGTATCGAAATCGGACTCGCCCTTCAAGCAATTCAGCATGGGAGGAGCACCA 826
OY 201 LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
Db 827 AAATCTGCTTATCTTAATGATGCAATGACACACACACACACACACACACAAATTAACCTA 886
OY 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGluArgGlnGlyThrGlnGlnMetLys 240
Db 887 GGCTAGAAAGCAAGCGATTAATGATTAACACAGAGAGATATCCAAAGAGATGAAG 946
OY 241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlnThrMetAspThrValAsnThrValMet 260
Db 947 GCTGCCGACAGAAAGTCTAAAGATCTGAAAGAACATGATCTGATCTGATCTGATG 1006
OY 261 IleAla 262
Db 1007 ATCGCG 1012

RESULT 8
ID AA159311 standard; DNA: 1048 BP.
AC AA159311;
XX
DT 22-APR-1997 (first entry)
DE DNA encoding C. pneumoniae 53 kDa antigen.
XX
KM antigen; antibody; detection; determination; epitope; ds.
XX Chlamydia pneumoniae.
OS
FH Key Location/Qualifiers
FT CDS 236..1012
FT /*tag= a
FT /note= "no stop codon present"


```

Db      811 CAACTCTGACATGACACCCCGGCGTCCCAACAAATATAGCTGTGCAACGAA 870
Qy      41 ThrlysglnileglnlnrtharaglnlylvsanThrsImetgluseraspAlaThr 60
      871 ACGAAGCAAAATACGCAACACGTCAGGTAATAACACTGAGATGGAACGATGCCACT 930
Qy      61 lIeAlaGlYAlaSerGlylYsAspLYsThrsSerThrThrlYsThrlGlnAlaPro 80
Db      931 ATTCTGTGCTCTGTGAAAAGACAAACTCTCGACTACAAAAACAGAAACAGCTCCA 990
Qy      81 GlnGlnGlyAlaAlaAlaGlyLysLysSerSerGlnSerGlnYsAlaGlyAlaAspThr 100
Db      991 CAACAGGAGAGTTCCTGCTGGGAAAGAAATCTCAGAAAAGTCAAAAGCGAGGTGCTGACT 1050
Qy      101 GlyAlaSerGlyAlaAlaAlaThrThraLaserasnThraThrLYsAlaMetGln 120
Db      1051 GGAGATACAGAGCGCGTCTACTACACATCAATACTGCACAAAATTTGCTATGCGAG 1110
Qy      121 ThrSerlIeGlnGlnAlaSerLYsSerMetGlnSerThrLeuGlnSerLeu 140
Db      1111 ACCCTATATGAGAGGCGCAAGCAAGTATGAGTCTACTTATGAGTCAACTCAAGCCTC 1170
Qy      141 SerAlaAlaGlnMetLysGlnValGlnAlaValAlaAlaAlaSerGlyLysSer 160
Db      1171 ACTGCCGCGCAAAATGAAAGATCGAAGCGGTTGTTGCTGCCCTCTCAAGGAAAGT 1230
Qy      161 SerGlySerAlaLysLeuGlnThrProGlnLeuProLYsProGlnYValThrProArgSer 180
Db      1231 TCGGGTTCGCAAAATGGAACACCTGAGCTCCCAAGCCCGGGGTGACACCAAGATCA 1290
Qy      181 GlnValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThr 200
Db      1291 GAGGTATGAAATGAGATCGAGTCCGGCTTCTTAAAGCAATTCAGACATGCGAAGCCACA 1350
Qy      201 LysSerAlaLeuSerasnThrAlaSerThrGlnAlaSerGlnThrAsnLYsLeu 220
Db      1351 AATTCGCTTATCTATCACTATGCAAGTACACCAAGCAAGCAACCAATTAACATA 1410
Qy      221 GlyLeuGlnLysGlnAlaIleLYsIleAspLYsGlnArgGlnGlnLYsGlnLeuLYs 240
Db      1411 GGTCTGAAAAGCAAGCAAGTAAATCGATTAAGAAAGCAAGAAATACCAAGATGAAG 1470
Qy      241 AlaAlaGlnGlnLysSerLYsAspLeuGlnGlyThrMetAspThrValAsnThrValMet 260
Db      1471 GGTGCCAAGCAAGATCTTAAAGATCTCGAAGCAACATGATCTGTCACTACTGTGATG 1530
Qy      261 lIeAlaValSerValAlaIleThrVal-----lIeSer 271
Db      1531 ATCCCGAA-GGGGTTCGAATTTGCCATGCGGGCCCTTAATTAACTCGAGAGATCCAG 1589
Qy      272 lIeValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeu-----AlaAlaGly 289
Db      1590 ATCTAATCATGATCTCTCTACGCGGACGATGTCGGCCGATCACCGGCGCACAGGT 1649
Qy      290 AlaAlaValGlyAlaAla----- 295
Db      1650 GCGGTTCCTGCGGCTTA-TATCCCGACATCACCGATGGGAGATCGGGCTCGCCACTT 1708
Qy      296 -----AlaAlaGlyGlyAla 300
Db      1709 CGGGCTCATGAGCGCTTTGTTGCGCGTGGATGATGTGGCAAGCCCGCGGGGAGCTG 1768
Qy      301 AlaGlyAla-----AlaAlaAla----- 306
Db      1769 TTGGGCGCCATCTCTTCATGACACCATTTCTTCTGGGCGGCGTCTCAACGCCCTCAAC 1828
Qy      306 ----- 306
Db      1829 CTACTACTGGGCTGCTCTCTAATGACAGAGTGCATAGAAGGAGAGCGTCGACCGATGCC 1888
Qy      307 -----ThThValAla 310

```

```

Db      1889 TTGACAGCCTTCAACCCAGTCACTCTTCGCGTGGCGCGGCGCATGACTATGTCGCC 1948
Qy      311 ThrGlnIleThrValGlnAlaValAlaGlnAlaValLYsGlnAlaValIleThrAlaVal 330
      1949 GCACATTATGACTGTCTCTTATCATGCAACTGTAAGACAG----- 1990
Qy      331 ArgGlnAlaIleThrAlaAlaIle-----LysAla 340
Db      1991 -----GTCCCGGCAAGCGCTCTGGGTCATTTCGGGAGAGACCGCTTTCGCTGAGC 2041
Qy      341 AlaValLysSerGlylIleLysAlaPheIleYsThrLeuValLYsAlaIleAlaLYsAla 360
Db      2042 GCGACGATGATCGCCTGCTGCGCTTGGGATTCGGAATCTTGACGCCCTCGCTCAAGCC 2101
Qy      361 lIeSerLYsGlylIleSerLYsValPheAlaLYsGlyThnGlnMetlIleAlaLYsAsnPhe 380
Db      2102 TTGCTACGTGTCGCC-----ACCAAGCTTTC 2131
Qy      381 ProLYsLeuSerLYsValIleSerSerLeu----- 390
Db      2132 GCGGAGAGACGAGCCATTCGCGCGCATGCGCGGCGGCGCGCTGAGCTTGTGCTG 2191
Qy      391 -----ThrSerLYsThrPValThrValGlyValGlyValAlaAlaAlaProAla 407
Db      2192 GCGTTCGCGAGCGGAGGCTGATGGCTTCCCATTAATGATTTCTTCGCTTCGCGGCG 2251
Qy      408 LeuGlyLYsGlylIleMetGlnMetGlnLeuSerGlnMetGlnGlnGlnAsnValAlaGlnPhe 427
Db      2252 ATCGGGATGCCCGCGGTTCGACGAGCATGCTGTCCAGTATGATTTCTTCGCTTCGCGG 2311
Qy      428 GlnLYsGlnValGlyLYsLeuGlnAlaAlaAlaAlaAspMetIleSerMet 443
      2312 CAGCTTCAA---GGATGCTCGCGGCTCTTACCAAGCTTACTTCGATC 2356
Db
RESULT 6
ID      AAT33893
AC      AAT33893;
XX      11-NOV-1996 (first entry)
DE      C. pneumoniae antigen (polypeptide A) expression vector.
KW      polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;
      PCPN533alpna; primer; assay; detection; antibody; diagnosis;
      infection; Chlamydia pneumoniae; ds.
XX      Synthetic.
OS      Synthetic.
FH      Key
FT      CDS
      Location/Qualifiers
      481..2428
      /*tag= a
W09609320-A1.
XX      28-MAR-1996.
PD      20-SEP-1995; 95WO-JP01896.
XX      28-APR-1995; 95JP-0106011.
PR      20-SEP-1994; 94JP-0224711.
XX      28-APR-1995; 95JP-0106006.
PR      28-APR-1995; 95JP-0106008.
XX      28-APR-1995; 95JP-0106009.
PR      28-APR-1995; 95JP-0106010.
XX      (HITB ) HITACHI CHEM CO LTD.
XX      Izutsu H, Matsumoto A, Obara K;
PI      WPI; 1996-188399/19.
XX      P-SDB; AAR94579.
DR

```

```

OY 164 AlAtysLeuGIuThrProGIuLeuProLysProGIyValThrProArgSerGIuValIle 183
DB 481 GCAAAATTTGGAACACGAGAGCTCCCAAGCCGGGGTGGACACCAAGATCAAGGTATAC 540
OY 184 GIUIEGlyLeuAlaLeuAlaLysAlaIleGIuThrLeuGIyGIuAlaThrLysSerAla 203
DB 541 GAAATCGGACCTGCGCTTCTTAACCAATTCAGACATTCGGAGAACCCCAAAATCTGCC 600
OY 204 LeuSerAntyrAlaSerThrGlnAlaGlnAlaAspGIuThrAspLysLeuGIu 223
DB 601 TTATCTAACTATGCAAGATACACACACACACCAACCAATTAATAGTCTAGCA 660
OY 224 LysGlnAlaIleLysIleAspLysGIuArgGIuGIyGlnIleuMetLysAlaIleGIu 243
DB 661 AAGCAAGCATTAATAATGATTAAGACAGAGAAAGATACCAAGATGAGAGCTGCCGA 720
OY 244 GlnLysSerLysAspLeuGIuGIyThrMetAspThrValAsnThrValMetIleAlaVal 263
DB 721 CAGAGTCTAAAGATCTCGAAGCAACAAATGATCTGCAATCTGATGATCGCGGTT 780
OY 264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGIyAlaGIuLeu 283
DB 781 TCTGTTCATTTACAGATTAATTTCTATTGTGCTGCTATTATTACATGCGGACCTGACTC 840
OY 284 AlaGIyLeuAlaAlaGIyAlaAlaValGIyAlaAlaAlaGIyGIyAlaAlaGIyAla 303
DB 841 GCTGGACATCGCTGCGCGAGAGCTGCTAGCTGCGAGCGGCGAGCTGAGAGAGCT 900
OY 304 AlaAlaAlaThrThrValAlaThrGlnIleThrValGIuAlaValGIuAlaValLys 323
DB 901 GCTGCCGCAACACGAGTACACACAAATTAAGTTCAGTTCAGTTCGCAACGCGGANA 960
OY 324 GlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaValLys 343
DB 961 CAAGCTGTATACAGCTGTACAGACAGGATACCGCGGCTATTAACCGGCTGTCANA 1020
OY 344 SerGIyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363
DB 1021 TCTGGAATTAAGCATTTATCAAACTTTAGTCAAAAGCATTTGCCAAACCATTTCTANA 1080
OY 364 GIyIleSerLysValPheAlaLysGIyThrGlnMetIleAlaLysAsnPheProLysLeu 383
DB 1081 GGAATCTTAAGCTTTTCGCTTAAGGAACTCAAAATGATTCGCAAGAACTTCCCAAGCTC 1140
OY 384 SerLysValIleSerSerLeuThrSerLysTrpValThrValGIyAlaGIyAlaVal 403
DB 1141 TCGAAAGTATCTCTGCTTACCAAGTAAATGGGTACGCGTGGGTTGGGTGTAGTT 1200
OY 404 AlaAlaProAlaLeuGIyLysGIyIleMetGlnMetGlnLeuSerGIuMetGlnIleAsn 423
DB 1201 GCGGCGCCCTCTCGGTAAAGGATATGCAAAATCGACCTCTCGAGATGCAACAAAC 1260
OY 424 ValAlaGlnPheGlnLysGIuValGIyLysLeuGlnAlaAlaAspMetIleSerMet 443
DB 1261 GTGCGTCAATTCAGAAACAAATCGGAAACACTGCAAGGCTGCGGTGATATGATTTCTANG 1320
OY 444 PheThrGlnPheTrpGlnGlnAlaSerLysIleAlaSerLysGlnThrGIyLysSerAsn 463
DB 1321 TTTCACCTCAATTTTGGCAACAGGCAAGTAAATTTGCCCTCAAAACCAAGGCAAGTCTAAT 1380
OY 464 GlnMetThrGlnLysAlaThrLysLeuGIyAlaGlnIleLeuLysAlaIleAlaIle 483
DB 1381 GAAATGACCTCAAAAGCTTCAAGCTGCGGCTCAAAATCTTAAGCGGTATGCGCGAATC 1440
OY 484 SerGIyAlaIleAlaGIyAla 490
DB 1441 AGCGGAGCATGCTGCGCA 1461

```

RESULT 5
 AAT14616
 ID AAT14616 standard; DNA; 5658 BP.
 XX
 AC AAT14616;

```

XX 07-NOV-1996 (first entry)
DT C. pneumoniae antigen (polypeptide A) expression vector pCPN533 alpha.
XX
DE Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;
XX pCPN533alpha; primer; assay; detection; antibody; diagnosis;
KW Infection; Chlamydia pneumoniae; ds.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 760..2647
FT /*tag= a
FT /*note= "STOP codon absent"
FT
XX
PN WO9609320-A1.
XX
PD 28-MAR-1996.
XX
PE 20-SEP-1995; 95WO-JP01896.
XX
PR 28-APR-1995; 95JP-0106011.
XX 20-SEP-1994; 94JP-0224711.
PR 28-APR-1995; 95JP-0106006.
PR 28-APR-1995; 95JP-0106008.
PR 28-APR-1995; 95JP-0106009.
PR 28-APR-1995; 95JP-0106010.
XX
PA (HITB ) HITACHI CHEM CO LTD.
XX
PI Izutsu H, Matsumoto A, Obara K;
XX
DR WPI: 1996-18839/19.
DR P-PSDB: AAR94579.
XX
PT Recombinant Chlamydia pneumoniae antigen and antibodies to it
PT used for detection and assay of C. pneumoniae e.g. in clinical
PT diagnosis
XX
PS Claim 13; Pages 77-84; 128bp; Japanese.
XX
CC The present sequence is an expression vector for the C. pneumoniae
CC polypeptide antigen, polypeptide A. C. pneumoniae strain YK41 was
CC cultured and genomic DNA extracted to prep. a lambda gt11 DNA
CC library. The library was then screened with an anti-YK41
CC monoclonal antibody (MAb), which was prepd. by fusing spleen cells
CC from a mouse infected with YK41 with myeloma P3/NS1/1-Ag4-1 to
CC produce a MAb expressing hybridoma. The DNA obtd. was then fused
CC with the expression vector pAD431 to give pCPN533alpha. The
CC plasmid was used to transform an E. coli host, which was cultured
CC to give the antigenic polypeptide, polypeptide A. Polypeptide A
CC and primers and probes derived from its DNA can be used in assays
CC for the detection of polypeptide A antibodies and DNA,
CC respectively, useful in the diagnosis of C. pneumoniae infection.
XX
SQ Sequence 5658 BP; 1446 A; 1481 C; 1424 G; 1307 T; 0 other;

```

Alignment Scores:
 Pred. No.: 9,8e-86 Length: 5658
 Score: 1282.00 Matches: 315
 Percent Similarity: 62.59% Conservative: 33
 Best Local Similarity: 56.65% Mismatches: 75
 Query Match: 54.12% Indels: 135
 DB: 17 Gaps: 10

US-09-889-314-2 (1-496) x AAT14616 (1-5658)
 OY 1 AspThrAsnMetSerIleSerSerSerGIyProAspAsnGlnLysAsnIleMetSer 20
 DB 751 GATACAAACATGCTAATTTCAATCTTACAGACCTGCAACAAATCAAAATATCATGCT 810
 OY 21 GlnValLeuThrSerThrProGlnGIyValProGlnGlnAspLysLeuSerGIyAsnGIu 40

Db 1384 GGTCCGCAACGACGCTAGCACACAAATTACAGTTCAAGCTGTGTCCAAAGCGGTGAAA 1443
 QY 324 GlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaValLys 343
 Db 1444 CAAGCTGTTATTCACAGCTGTGACAGCAAGCATACCGCGGCTATTAAGCGGTGTGCAAA 1503
 QY 344 SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363
 Db 1504 TCGGATATAAAGCATTTATCAAAACTTTAGTCAAAAGCATTTGCCAAAGCCATTTCTAAA 1563
 QY 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeu 383
 Db 1564 GGATCTCTAAGGTTTGGCTAAGGAGATCAATGATGCGAGAACTTCCCAAGCTC 1623
 QY 384 SerLysValIleSerSerLeuThrSerLysTrpValThrValGlyValGlyValVal 403
 Db 1624 TCGAAGTCTATCTGCTCTTACAGTAATGGGTGACGGTTGGGGTTGGATTAGTT 1683
 QY 404 AlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnLeuGlnAsn 423
 Db 1684 GCGGCGCTCTCGGTAAAGGATTTATGCAATGACGCTTCGGAGATGCAACAAAC 1743
 QY 424 ValAlaGlnPheGlnLysGlyValGlyLysLeuGlnAlaIleAlaAspMetIleSerMet 443
 Db 1744 GTGCTCAATTTTCAGAAAGAGTCGAAACCTGCAGGCTCGGCTGATGATTTCTATG 1803
 QY 444 PheThrGlnPheTrpGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyGlnSerAsn 463
 Db 1804 TTGCTCAATTTTGGCAACAGGCAAGTAATTTGCTCAAAACAAACAGCGAGTCTAAT 1863
 QY 464 GluMetThrGlnLysAlaThrLysLeuGlyValGlnIleLeuLysAlaThrAlaIle 483
 Db 1864 GAAATGACTCAAAAAGCTACCAAGCTGGCGCTCAAAATCTTAAAGCTATGCGGCATC 1923
 QY 484 SerGlyAlaIleAlaGlyAla 490
 Db 1924 AGCGAGCATCGCTGGCGCA 1944

RESULT 4
 AAT14612
 ID AAT14612 standard; DNA: 1464 BP.
 AC AAT14612;
 XX
 XX 07-NOV-1996 (first entry)
 DE Chlamydia pneumoniae polypeptide antigen (polypeptide A) DNA.
 KW Polypeptide antigen; polypeptide A; strain YK41; plasmid; probe;
 KW PCN533alpha; primer; assay; detection; antibody; diagnosis;
 KW Infection; ds.
 XX Chlamydia pneumoniae.
 OS
 XX
 FH Key 1..1464
 FT CDS Location/Qualifiers
 FT 1..777 /tag= a
 FT misc_feature /note= "STOP codon absent"
 FT /tag= b
 FT /note= "claimed 5'-end fragment"

WO9609320-A1.
 XX
 PD 28-MAR-1996.
 PD 20-SEP-1995; 95WO-JP01896.
 PF
 XX 28-APR-1995; 95JP-0106011.
 PR 20-SEP-1994; 94JP-0224711.
 PR 28-APR-1995; 95JP-0106006.
 PR 28-APR-1995; 95JP-0106008.
 PR 28-APR-1995; 95JP-0106009.

PR 28-APR-1995; 95JP-0106010.
 XX
 XX (HTB) HITACHI CHEM CO LTD.
 XX
 PI Izutsu H, Matsumoto A, Obara K;
 XX
 DR MPI: 1996-188399/19.
 DR P-PSDB: AAR94579.
 XX
 PT Recombinant Chlamydia pneumoniae antigen and antibodies to it
 PT used for detection and assay of C. pneumoniae e.g. In clinical
 PT diagnosis
 PS Claim 9; pages 60-64; 128pp; Japanese.
 XX
 CC The present sequence encodes the C. pneumoniae polypeptide antigen,
 CC polypeptide A. C. pneumoniae strain YK41 was cultured and genomic
 CC DNA extracted to prep. a lambda gt11 DNA library. The library was
 CC then screened with an anti-YK41 monoclonal antibody (MAb), which
 CC was prep. by fusing spleen cells from a mouse infected with YK41
 CC with myeloma F3/NS1/-Ag4-1 to produce a MAb expressing hybridoma.
 CC The DNA obcd. was then fused with the expression vector pAD4431 to
 CC give pcPN533alpha. The plasmid was used to transform an E. coli
 CC host, which was cultured to give the antigenic polypeptide, from
 CC polypeptide A. Polypeptide A and primers and probes derived from
 CC its DNA can be used in assays for the detection of polypeptide A
 CC antibodies and DNA, respectively, useful in the diagnosis of
 CC C. pneumoniae infection.
 XX
 SQ Sequence 1464 BP; 470 A; 333 C; 346 G; 315 T; 0 other:
 XX

Alignment Scores:
 Pred. No.: 3,65e-162 Length: 1464
 Score: 2300.00 Matches: 485
 Percent Similarity: 99.598 Conservative: 0
 Best Local Similarity: 99.598 Mismatches: 2
 Query Match: 97.098 Indels: 0
 DB: 17 Gaps: 0

US-09-889-314-2 (1-496) x AAT14612 (1-1464)
 QY 4 MetSerIleSerSerSerSerGlyProAspGlnLysAsnIleMetSerGlnValLeu 23
 Db 1 ATGCTATTTCACTCTCTCAGACCTGCACAAATCAAAATATCATGTCAGTTCTG 60
 QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyLysGln 43
 Db 61 ACATCGACACCCAGGCGCGTCCCAACAGATAGCTGTCTGCAACGAAGCAAGCA 120
 QY 44 IleGlnIleThrArgGlnGlyLysAsnThrGluMetGlnSerAspAlaThrIleAlaGly 63
 Db 121 ATACAGCAACACGTCAGGTTAAACACAGATGAGTGAAGCATGCGACTATTTGCTGT 180
 QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83
 Db 181 GCTTGTGAAAAGACAAACTCTCTCGACTACCAAAAACAAACAGCTCCACAAAGGCA 240
 QY 84 ValAlaIleAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThrGlyValSer 103
 Db 241 GTTGTCTGTGGCAAGATCTCTCGAAGATCAAAAGCAGAGTGTGATCTGAGATCA 300
 QY 104 GlyAlaAlaIleThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
 Db 301 GGAGCGGCTCTACTACAGCATCAAAATATGCAAAATATGCTATGACAGACTCTATT 360
 QY 124 GlnGluAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143
 Db 361 GAAAGGCGGACCAAAAGATGATGACTTACCTTAAGTACACTTCAAAACCTCAGCGCGC 420
 QY 144 GlnMetLysGluValGluAlaValAlaValAlaIleLeuSerGlyLysSerSerGlySer 163
 Db 421 CAATGAAAGAGTCGAAGCGGTTGTGCTGCTGCCCTCTCAGGAGAAAGTTGCGGTTCC 480

QY 461 GUSERASGLUMETThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyr 480
DB 926540 GAGCTATGTAATGATCATCAAAAGCTACCAAGCTGGCGCTCAAAATCCTTAAAGCGTAT 926481
QY 461 AlaAlaIleSerGlyAlaAlaIleAlaGlyAlaHisLysThrAsnAsnPhe 496
DB 926480 GCCGCAATCAGCGGAGCCATCGTGGCGAGCATMAAACCAATAATTTT 926433

RESULT 3
AAT14618
ID AAT14618 standard; DNA; 1947 BP.
AC AAT14618;
XX
DT 11-NOV-1996 (first entry)
DE DHFR/C. pneumoniae antigen fusion protein (polypeptide B) DNA.
XX
KW Polypeptide antigen; polypeptide B; strain YK41; plasmid; probe;
KW PCPN533T; primer; assay; detection; antibody; diagnosis;
KW infection; fusion protein; dihydrofolate reductase; DHFR;
KW Chlamydia pneumoniae; ds.
XX
OS Synthetic.
XX
FH Key Location/Qualifiers
FT CDS 1..1947
FT /*tag= a
FT /note= "STOP codon absent"
XX
PN MO9609320-A1.
XX
PD 28-MAR-1996.
XX
PE 20-SEP-1995; 95WO-JP01896.
XX
PR 28-APR-1995; 95JP-0106011.
PR 20-SEP-1994; 94JP-0224711.
PR 28-APR-1995; 95JP-0106006.
PR 28-APR-1995; 95JP-0106008.
PR 28-APR-1995; 95JP-0106009.
PR 28-APR-1995; 95JP-0106010.
XX
PA (HITB) HITACHI CHEM CO LTD.
XX
PI Izutsu H, Matsumoto A, Obara K;
XX
DR WPI; 1996-188399/19.
DR P-PSDB; AAR94584.
XX
PT Recombinant Chlamydia pneumoniae antigen and antibodies to it -
PT used for detection and assay of C. pneumoniae e.g. in clinical
PT diagnosis
XX
XX Claim 25; Pages 97-103; 128pp; Japanese.
XX
XX The present sequence encodes the dihydrofolate reductase (DHFR)/
CC C. pneumoniae polypeptide antigen fusion protein, polypeptide B.
CC C. pneumoniae strain YK41 was cultured and genomic DNA extracted
CC to prep. a lambda gt11 DNA library. The library was then screened
CC with an anti-YK41 monoclonal antibody (Mab), which was prepd. by
CC fusing spleen cells from a mouse infected with YK41 with myeloma
CC P3/NS1/1-7g4-1 to produce a Mab expressing hybridoma. The DNA
CC obtd. was then fused with DHFR DNA and the expression vector
CC pAD431 to give PCPN533T. The plasmid was used to transform an
CC E. coli host, which was cultured to give the antigenic polypeptide
CC fusion protein, polypeptide B. Polypeptide B and primers and
CC probes derived from its DNA can be used in assays for the
CC detection of antigenic polypeptide antibodies and DNA,
CC respectively, useful in the diagnosis of C. pneumoniae infection.
XX
SQ Sequence 1947 BP; 591 A; 452 C; 483 G; 421 T; 0 other;

Alignment Scores:
Pred. No.: 9,24e-163 Length: 1947
Score: 2310.00 Matches: 485
Percent Similarity: 100.00%
Best Local Similarity: 99.59% Mismatches: 2
Query Match: 97.51% Indels: 0
DB: 17 Gaps: 0

US-09-889-314-2 (1-496) x AAT14618 (1-1947)

QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValIleu 23
DB 484 ATGCTATTTCATCTTCTTCAGAGACCTGACATCAAAAAATATCATGCTCAAGTTCG 543
QY 24 ThrSerThrProGlnGlyValProGlnAspLysLeuSerGlyAsnGluThrLysGln 43
DB 544 ACATCGACACCCCGAGCGCTGCCCAACAGATAGCTGCTGGCAACGAAAGCA 603
QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGluMetGluSerAspAlaThrIleAlaGly 63
DB 604 ATACAGCAACACGTCAGGGTAAACACTGAGATGAAAGGATGCCACTATTGCTGGT 663
QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGly 83
DB 664 GCTTCGAAAGACAAACCTTCCTGACTACAAAACAGAAACAGCTCCACACAGCGGA 723
QY 84 ValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyValSer 103
DB 724 GTTGCTGCTGGGAAAGAAATCTCAGAAAGTCAAAAGCGTGTGATCTGAGATATCA 783
QY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
DB 784 GAGCGGCTGCTACTACACACTCAAAATCTGACACAAAATTTGCTATGCACTCTATT 843
QY 124 GluGluAlaSerLysSerMetGluSerThrLysLeuGlnSerLeuSerLeuSerAlaAla 143
DB 844 GAAGAGCGCAGCAAAAGTATGAGTCTACTTACATCAAGTCAAAAGCCTCAGTGGCGG 903
QY 144 GlnMetLysGluValGlnAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlySer 163
DB 904 CAATGAAAGAAAGTCAACCGGTTGTTGTTGCTCCCTCAGGGGAAAGTTCCGGTTCC 963
QY 164 AlaLysLeuGluThrProGluLeuProLysProGlyValThrProArgSerGluValIle 183
DB 964 GCAAAATTGGAAACACCTACCTCCCCAAGCCCGGGGTACACCAAGATCAAGGTTATTC 1023
QY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThrLysSerAla 203
DB 1024 GAATCGGACCTCGCGCTGCTTAAGCAATTCAGACATTGGAGAAAGCCAAATCTGCC 1083
QY 204 LeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlu 223
DB 1084 TTATCTTAACCTATGCAAGTCAACAGCACAGCAAGCAAAATTAATCTAGTCTAGAA 1143
QY 224 LysGlnAlaIleLysIleAspLysGluArgGluGluTyrGlnGluMetLysAlaIleGlu 243
DB 1144 AAGCAAGCGATTAATAATCATTAAGACAGAAAGATTAACAAGATGAAGCTGCGGAA 1203
QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaVal 263
DB 1204 CAGAAGTCAAAAGATCTCAAGAGCAACAAATGATGTCATATGATGATCGGGTT 1263
QY 264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu 283
DB 1264 TCTGTGCCATTACAGTATTTCTATTGTTGCTGCTATTTTACATGCGGACCTGACAC 1323
QY 284 AlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaAlaAlaGlyAlaAlaGlyAla 303
DB 1324 GCTGCACTGCTGCGGAGCTGCTGTAGTGCAGCGGCAAGCTGGAAGTCCACAGAGAGCT 1383
QY 304 AlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLys 323

XX AAX91990;
 AC 13-SEP-1999 (first entry)
 DT Nucleotide sequence of the complete genome of Chlamydia pneumoniae.
 XX Respiratory disease; pneumonia; bronchitis; heart disease; sarcoidosis;
 XX sinusitis; purulent otitis media; erythema nodosum; pharyngitis;
 KM vaccine; neutralising epitope; ss.
 XX Chlamydia pneumoniae.
 OS
 XX W09927105-A2.
 XX 03-JUN-1999.
 PD 20-NOV-1998; 98MO-IB01890.
 XX 04-NOV-1998; 98OS-0107078.
 PR 21-NOV-1997; 97FR-0014673.
 XX (GEST) GENSET.
 PA Griffais R;
 PI WPI; 1999-357842/30.
 DR Genome sequence of Chlamydia pneumoniae
 XX Claim 1; Page 291-611; 1912pp: English.
 PS The present sequence represents the complete genome of Chlamydia
 XX pneumoniae, and encodes proteins AAY34584-Y35879. C. pneumoniae causes
 CC respiratory disease such as pneumonia and bronchitis, and is thought
 CC to be a contributing factor in heart disease, sarcoidosis, sinusitis,
 CC purulent otitis media, erythema nodosum or pharyngitis. The polypeptides
 CC encoded by the open reading frames of the C. pneumoniae genome (see
 CC AAY34584-Y35879) can be used in immunogenic compositions as vaccines.
 CC Vectors containing C. pneumoniae nucleotide sequences can also be
 CC used as immunogenic compositions, especially where the vector directs
 CC the expression of a neutralising epitope of C. pneumoniae.
 XX
 SQ Sequence 1230025 BP; 367213 A; 249833 C; 249013 G; 363589 T; 377 other;
 Alignment Scores:
 Pred. No.: 1.02e-162 Length: 1230025
 Score: 2354.00 Matches: 493
 Percent Similarity: 99.40% Conservative: 0
 Best Local Similarity: 99.40% Mismatches: 3
 Query Match: 99.37% Indels: 0
 DB: Gaps: 0
 US-09-889-314-2 (1-496) x AAX91990 (1-1230025)
 OY 1 AsptRanMetSerIleSerSerSerGlyProAspAnGlnLysAsnIleMetSer 20
 DB 927920 GATACAAACATCTCTATTTCATCTCTTCAGAGACTGCACAAATAATATCTGCT 927861
 OY 21 GlnValIleuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlu 40
 DB 927860 CAAGTTCGACATCGACACCCAGGCGTCCCAACAAAGATTAAGCTGCTGCAACGAA 927801
 OY 41 ThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGluMetGluSerAspIleThr 60
 DB 927800 ACGAAGCAAAATACACAAACAGTCAAGGTAATAAACACATGAGATGGAAGCGATCCACT 927741
 OY 61 IleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaPro 80
 DB 927740 ATTGCTGGTCTTCTGGAAGAAACAAACTCTCTGACGTACAAAAACGAAACAGCTCA 927681
 OY 81 GlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThr 100

162 (2) m 2003

DB 927680 CAACAGGAGTTGCTGCGGAAAGATCTCGAAAGTCAAAAGCAGCGTCTGATACT 927621
 OY 101 GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120
 DB 927620 GGAGTATCAGAGAGCGGCTCTCTACTACACATCAAAATACGCAAAATAATCTATGCG 927561
 OY 121 ThrSerIleGlnGlnLysSerLysSerMetGluSerThrLeuGlnLysLeuSerLeu 140
 DB 927560 ACCCTATTGAAAGAGCGGACCAAAATATGAGACTCTACTTGAAGCACTTCAAGGCCCTC 927501
 OY 141 SerAlaAlaGlnMetLysGluValAlaValAlaValAlaLeuSerGlyLysSer 160
 DB 927500 AGTCCGCGCAATTAAGAAAGTCGAAGCGGTGTGTTGCTGCTCCCTCAGGGAATACT 927441
 OY 161 SerGlySerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSer 180
 DB 927440 TCGGTTCCCGCAAAATTTGGAACACCTGAGCTCCCAACCGCGGGGTGACCCAAATCA 927381
 OY 181 GluValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnLysAlaThr 200
 DB 927380 GAGCTTATCGAAATTCGAGCTCGCGCTTCTGTAAGCAATTCAGACATTGGGAGAGCCACA 927321
 OY 201 LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
 DB 927320 AAATCTGCTTATCTAATCTATGCAAGTACCAAGCAACCAACCAACCAATTAACGA 927261
 OY 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGluArgGlnGlnThrGlnGlnMetLys 240
 DB 927260 GGTCTAGAAAAGCAAGGATTAATAATGATTAAGAACGAAACAAATACCAAGATGAAG 927201
 OY 241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMet 260
 DB 927200 GCTCCGACACAGAAAGTCTAAAGATCTCGAAGGAACATGATCTGCTGATGATG 927141
 OY 261 IleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrLysGly 280
 DB 927140 ATCCGGTTCTGTTGTCATACAGTTATTTCTATGTTCTGCTATTTTACATGCGGA 927081
 OY 281 AlaGlyLeuAlaGlyLeuAlaAlaAlaGlyAlaAlaValGlyAlaAlaAlaAlaGlyAla 300
 DB 927080 GCTGAGCTCGTGCAGCTCGCTGCGGAGCGTCTGAGTGCAGCGGAGCTGGAGGTGCA 927021
 OY 301 AlaGlyAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaGln 320
 DB 927020 GCAGAGAGCTGCTCCGCAACACACGCTGACCAACAAATACAGTTACAGTGTGTCCAA 926961
 OY 321 AlaValLysGlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAla 340
 DB 926960 GCGGTGAACAAAGCTGTTATACAGCTGTGCAGACAGCATCACCGCGCTATTAAGCG 926901
 OY 341 AlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAla 360
 DB 926900 GCTGCAAAATCTGGAATTAAGCATTTATCAAACTTATGCAAAAGCATGCGCAAAAGCC 926841
 OY 361 IleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAspPhe 380
 DB 926840 ATTCTTAAGCAATCTTAAGGTTTTCGTAAAGCAACTCAAAAGATTCGAAACACTTC 926781
 OY 381 ProLysLeuSerLysValIleSerSerLeuThrSerLysThrValThrValGlyValGly 400
 DB 926780 CCCAAGCTCTGGAAGTATCTCGTCTTACAGTAATAGGTCAGCGTGGGGTGGG 926721
 OY 401 ValValValAlaAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMet 420
 DB 926720 GTTGTAGTTGCGCGCTGCTGCTCGTAAAGGATTAATGCAAAATGCAAGCTTCGGAATG 926661
 OY 421 GlnGlnAsnValAlaGlnPheGlnLysGluValGlyLysLeuGlnAlaAlaAlaAspMet 440
 DB 926660 CAACAAAACGTCGCAATTCAGAAAGAGTCGAAAGATGCAAGCTCGCGCTGATATG 926601
 OY 441 IleSerMetPheThrGlnPheThrProlGlnAlaSerLysIleAlaSerLysGlnThrGly 460
 DB 926600 ATTCTATGTTCACTCAATTTTGGCAACAGCAAGTAAATTTGCCCTCAAAACAAACAGCG 926541

```
XX MO200046359-A2.
XX 10-AUG-2000.
XX 28-JAN-2000; 2000WO-GB00237.
XX 05-FEB-1999; 99GB-0002555.
XX (NEUT-) NEUTEC PHARMA PLC.
XX Burnie JP, Matthews RC;
XX WPI: 2000-543485/49.
XX P-PSDB: AAB08322.
XX New Chlamydia pneumoniae protein of 496 amino acids for diagnosing,
XX preventing and treating C. pneumoniae infection and atherosclerosis,
XX including coronary atherosclerosis -
XX Claim 3; Page 27-29; 35pp; English.
XX The present sequence encodes a Chlamydia pneumoniae protein. The
XX protein, immunogenic fragments of it, nucleotide sequences encoding
XX it, or inhibitor specific agents it are used to manufacturing
XX a medicament for the treatment of infection due to C. pneumoniae.
XX An antibody specific against the protein can diagnose a C. pneumoniae
XX infection. C. pneumoniae infection can be prevented. Atherosclerosis,
XX including coronary atherosclerosis, caused by C. pneumoniae can also
XX be prevented or treated.
XX
XX Sequence 1491 BP; 485 A; 335 C; 348 G; 323 T; 0 other:
XX
XX Alignment Scores:
XX Pred. No.: 2,75e-167 Length: 1491
XX Score: 2369.00 Matches: 496
XX Percent Similarity: 100.00% Conservative: 0
XX Best Local Similarity: 100.00% Mismatches: 0
XX Query Match: 100.00% Indels: 0
XX DB: 21 Gaps: 0
XX
XX US-09-889-314-2 (1-496) x AAA63621 (1-1491)
OY 1 AsptRrAsnMeSerTlSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSer 20
DB 1 GATCAAAACATCTCTATTTCATCTTTCAGACCTGCACAAATCAAAAATATTCATGTCT 60
OY 21 GlnValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGln 40
DB 61 CAAGTTCTGACATCGACACCCAGGGCGTGGCCCAACAAGATTAAGCTGTCTGCAACGAA 120
OY 41 ThrLysGlnIleGlnGlnThrArgGlnGlnLysAsnThrGluMetGluSerAspAlaThr 60
DB 121 ACGAAGCAAAATACACCAAAACGTCAGGGCTAAACACATGATGGAAAGCCATGCTTCT 180
OY 61 IleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaPro 80
DB 181 ATTGCTGTGCTTCTTGGAAGAACAAACTTCTCGACTACAAAACACAAACAGCTCCA 240
OY 81 GlnGlnGlyAlaIleAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThr 100
DB 241 CAACAGGGAGGTTGCTGCTGGGAAAGATCTCAGAAAGCAAAAGCAGAGCTGATACT 300
OY 101 GlyValSerGlyAlaIleAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGln 120
DB 301 GGAGTATACAGGCGGCTGCTACTACAGCATCAAAATACTGCAACAAATTCCTATGCGAG 360
OY 121 ThrSerIleGlnGlnLysSerLysSerMetGluSerThrLeuGlnSerLeuGlnInsleu 140
DB 361 ACCTCTATTGAAGAGCGGCAAAAGATGAGTCTACCTTAGAGTCACTTCAAAAGCCCTC 420
OY 141 SerAlaAlaGlnMetLysGluValGlnAlaValAlaIleAlaLeuSerGlyLysSer 160
DB 141 SerAlaAlaGlnMetLysGluValGlnAlaValAlaIleAlaLeuSerGlyLysSer 160
```

```
DB 421 AGTCCCGCAAAATGAAGATCGAAGCGGTTGTGTGCTCCCTCTCAGGAAAAAGT 480
OY 161 SerGlySerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSer 180
DB 481 TCGGGTCCGCAAAATGGAACACCTGAGCTCCCAAGCCGGGGGTGACACCAAGATCA 540
OY 181 GlnValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlyAlaThr 200
DB 541 CAGGTTATCGAAATGCGACTCCGCTTGTCTAAGCAATTCACACTTGGAGAACCCAGA 600
OY 201 LysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
DB 601 AAATCGCTTATCTAATCATGCAAGTACACCAAGCAAGCAGCAACCAAAATTAACCA 660
OY 221 GlyLeuGlnLysGlnAlaIleLysIleAspLysGluValGlnGlnThrGlnIleMetLys 720
DB 661 GGTCTGAAAGCAAGCCGATTAATAATCGATTAAGAAGCAAGAAATACCAAGAGATGAG 780
OY 241 AlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMet 260
DB 721 GCTGCCGAACAGAACTCTAAGATCTCGAAGAACAAATGATACTGTCATCTGTGATG 780
OY 261 IleAlaValSerValAlaIleThrValIleSerIleValAlaIleIleThrCysGly 280
DB 781 ATCGGGGTTCTGTGCTGATACATGATATCTAATGTTGCTGATATTTTACATGCGGA 840
OY 281 AlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAlaAlaIleGlyAla 300
DB 841 GCTGACCTCGGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 900
OY 301 AlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 320
DB 901 GCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 960
OY 321 AlaValLysGlnAlaValIleThrAlaValAlaValAlaIleThrAlaAlaIleLysAla 340
DB 961 GCGGTGAAGCAAGCTGTTATCAGCTGTCAGCAAGCATCAAGCATCAAGCATCAAGCAT 1020
OY 341 AlaValLysSerGlyIleLysAlaIleLysAlaIleLysAlaIleLysAlaIleLysAla 360
DB 1021 GCTGTCAAAATGGAATTAACCAATTAATTAATTAATTAATTAATTAATTAATTAAT 1080
OY 361 IleSerLysGlyIleSerLysValPheAlaLysGlyThrGluMetIleAlaLysAsn 380
DB 1081 ATTCTGAAGCAATCTCTAAGCTTTTCTGTAAGGCAACCAATGATTTCCGAAGAATCT 1140
OY 381 ProLysLeuSerLysValIleSerSerLeuThrSerLysThrValThrValGlyAla 400
DB 1141 CCCAAGCTCTCGAAAGTCAATCTGCTCTTACCAAGTAATGGGTCCAGGTGGGTGGG 1200
OY 401 ValValValAlaAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGluMet 420
DB 1201 GTTGTAGTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
OY 421 GlnGlnAsnValAlaGlnPheGlnLysGlnValGlyLysLeuGlnAlaIleAlaAspMet 440
DB 1261 CAACAAAGCTGCTCAATTCAGAAAGAAAGTCGGAAGAACTCGAGGCTCGGCTGATAG 1320
OY 441 IleSerMetPheThrGlnPheThrPheGlnAlaSerLysIleAlaSerLysGlnThrGly 460
DB 1321 ATTTCTATGTTCAATCAATTTTGGCAACAGGCAAGTAATTAATTTGCCCAAAACAGCG 1380
OY 461 GluSerAsnGluMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaThr 480
DB 1381 GAGTCTAAAGAAATGATCAAAAAAGCTACCAAGCTGAGGCTCAATCTTTAAAGCTAT 1440
OY 481 AlaAlaIleSerGlyAlaIleAlaIleAlaIleLysThrAsnAsnPhe 496
DB 1441 GCCGCATACAGGGAAGCCATCGCTGGCGCACATTAATTAATTTT 1408
RESULT 2
AAK91990/C
ID AAK91990 standard; DNA; 1230025 bp.
```



```
OY 400 GlyValValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlu 419
      ::::: ||||| :::: |||
Db 3202 AATATCATTTCAAAA-----AACGGTATTAACACCGTACTGTTAAAGGC 3246
      ::::: ||||| :::: |||
OY 420 MetGlnGlnAsnValAlaGlnPheGlnLysGluValGlyLysLeuGlnAlaAlaAsp 439
      ::::: ||||| :::: |||
Db 3247 GTTAAATTTGATGTAATACATTCACCGCGGTATA-----GCAAGCGTAGAT 3294
      ::::: ||||| :::: |||
OY 440 MetIleSerMetPheThrGlnPheThrGlnGlnAlaSerLysIleAlaSerLysGlnThr 459
      ::::: ||||| :::: |||
Db 3295 GAAGTAATT-----GAAGCGAAACGCATCCTTGAGAGGTTAAAA 3333
      ::::: ||||| :::: |||
OY 460 GlyGlnSerAsnGlnMetThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAla 479
      ::::: ||||| :::: |||
Db 3334 GATTATCTGATGAAGAAGACGCTTAGCTAAACTTGGCGTAAGCGCTGTACGT--- 3390
      ::::: ||||| :::: |||
OY 480 TyrAlaAlaIleSerGlyAlaIle 487
      ::::: ||||| :::: |||
Db 3391 TTTGCTGAGCCAAATATATGCCATT 3414
      ::::: ||||| :::: |||
```

Search completed: January 27, 2003, 16:22:19
Job time : 3137 secs

```

QY 341 a-----VallySerGlylleYsaAlaPheIleYsThrLeuVallySaIleI 357
|          |||          |||          |||          |||
Db 2943 CTTACCCCAAGAGCTCTCGCGCCCGCCAGCTGTGTGCAGAGCTGCAAGCAGT 3002

QY 357 eaIaYsaAlaIleSerlySlyIleSerlySaIlePheAlaIleGlyGlyThrGlnIleAl 377
|          |||          |||          |||          |||
Db 3003 GGCAGAGCAAGATTCCAGCTGTGTGCAGAGGCGTC-----CGAGGAAGCCAA--GCCCA 3053

QY 377 alyAsnPheProlySleuSerlySaIleSerSerIleYsThrPValThrVa 397
|          |||          |||          |||          |||
Db 3054 GCCTGACAGCCCAAGCCCTAGCTTGCCTCATTTGCTGCCAGCAGAGCTTCCTGCCAGCC 3113

QY 397 lGlyValGlyValValAlaIle-----ProAlaLeuGlylySgIlyIleAl 413
|          |||          |||          |||          |||
Db 3114 AGGTGGAGAGATGGTGTGCAGCTGCAAGAGCCCTCAGTGCACAGATTCCAGGACCAAGCTTC 3173

QY 413 tGlnMetGlnLeuSerGlnMetGlnGlnAn-----ValAlaGlnPheGlnIly 429
|          |||          |||          |||          |||
Db 3174 AGCCATCCAGCTAGTCAAGTGTGCCAAGAACCTGGGCAAGCCGCGCTGCTGAACCTCGGAC 3233

QY 429 sGlyValGlyLySleuGlnAlaAlaIleAspMetIleSerMet 443
|          |||          |||          |||          |||
Db 3234 GCGTGGCCCAAGAGCTCAGAGAGCATGTGCACCTTTGGAGATG 3276

RESULT 45
US-09-206-942-64
; Sequence 64, Application US/09206942
; Patent No. 6432669
; GENERAL INFORMATION:
; APPLICANT: Loosmore, Sheena M.
; APPLICANT: Yang, Yan-Ping
; APPLICANT: Klein, Michel H.
; TITLE OF INVENTION: Protective Recombinant Haemophilus Influenzae High
; FILE REFERENCE: 1038-861 MIS:jb
; CURRENT APPLICATION NUMBER: US/09/206,942
; CURRENT FILING DATE: 1998-12-08
; EARLIER APPLICATION NUMBER: 09/167,568
; EARLIER FILING DATE: 1998-10-07
; NUMBER OF SEQ ID NOS: 95
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 3543
; TYPE: DNA
; ORGANISM: Haemophilus influenzae
; US-09-206-942-64

Alignment Scores:
Pred. No.: 9.75e-05 Length: 3543
Score: 146.50 Matches: 114
Percent Similarity: 36.55% Conservative: 79
Best Local Similarity: 21.59% Mismatches: 228
Query Match: 6.18% Indels: 107
DB: 4 Gaps: 23

US-09-889-314-2 (1-496) x US-09-206-942-64 (1-3543)

QY 7 SerSerSerSerGlyProAspAsnGlnIlySaAlIleMetSerGlnValIleuThrSerThr 26
|          |||          |||          |||          |||
Db 2011 AACACTGAAGTGGCAGAGTGACAAACAATATCCGCT-----TTAATCTTATCT 2055

QY 27 ProGlnGlyValProGlnGlnAspIlySleuSerGlyAsnGlnIlyThrLySglnIleGln 46
|          |||          |||          |||          |||
Db 2056 GCAAAAGATGTAGAAATAACAACATATTTACTTCTACAAAAAGTGAAGCTCTCTGCG 2115

QY 47 ThrATGlnGlyLySaAsnThrGlnMetGlnSerAspAlaThrIleAlaGlyValaSerGly 66
|          |||          |||          |||          |||
Db 2116 GCAAAATGAGGATTAACCT-----AAACAGGTACACCATTAATCCACCGCGCT 2169

QY 67 LysAspLySThrSerSerThrThrIlySThrGlnIlyThrAlaProGlnGlnIlyValAlaIle 86
|          |||          |||          |||          |||
Db 2170 AACGTGGAGATTAACCGCT-----CATACAGCAGCATATCCAAAGCGCAATTGAG 2217

```

```

QY 87 GlyLyGlnSerSerGlnSerGlnIlySaIleAlaIleAspThr-----GlyVal 102
|          |||          |||          |||          |||
Db 2218 TCCAGCTGCGCTGTGTGCAATTTGTGGCAGCGCGCATACTTGTGCTAGTAAATAT 2277

QY 103 SerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrIlyIleAlaMetGlnThrSer 122
|          |||          |||          |||          |||
Db 2278 TCAGGCAAGCGCCTTACTCTTACTGCAATATACGGT-----GCATTAACCACTTTG 2328

QY 123 lIleGlnAlaIleSerlySleuSerMetGlnSerThrLeuGlnSerLeuGlnSerIleuSerAla 142
|          |||          |||          |||          |||
Db 2329 GCAGGCTTACAATTAAGAAACGAGAGTAAACCACTTCAAGTCAATCAAGTATATATC 2388

QY 143 AlaGlnMetLySglnValGlnAlaValAlaIleAlaIleuSerGly---LysSerSer 161
|          |||          |||          |||          |||
Db 2389 GCGGTAATAATTTCCGCGCAAGACAGTAAACGTTAAAGCAACATAATGTTAACCAACCA 2448

QY 162 GlySerAlaLySleuGlnIlyThrProGlnLeuProLySProGlyValThrProArGserGln 181
|          |||          |||          |||          |||
Db 2449 GCAGACTCAAAATTTGAGGAGACTGAA---GGCGAGCTAATGTACAAGCAAAACAAAGC 2505

QY 182 ValIle-----GlnIleGlyLeuAlaIleuAlaIleGlnIlyThrLeuGlyGlnAla 199
|          |||          |||          |||          |||
Db 2506 ATATATTGCGGTCATAATTTCTGTGTGCACAGTAAAGTTAACCGCGCAGAGGTTTAAAC 2565

QY 200 ThrLySerAlaIleuSerAsnTyraIleSerThrGlnAla-----GlnAlaAspGlnThr 217
|          |||          |||          |||          |||
Db 2566 ACCCAAGCAGGC---TCTACGATTACTGGACCGAGACGCTGACCACTTCAAGCCAAATCA 2622

QY 218 AsnLySleuGly-----LeuGlnLySglnAlaIleIys----- 228
|          |||          |||          |||          |||
Db 2623 GGTAAATTCGCGGCATCATTTCTGTGTGCAGCAAGTAAATTAAGCCCAACCAAGATTTA 2682

QY 229 lIleAspLySglnValGlnIlyThrGlnIlyMetLySAlaIleGlnIlySleuSerLySaP 248
|          |||          |||          |||          |||
Db 2683 ATTAATTAATTCGCGTTCAGAGATT---AAAGCAAGCGCGCGCAGGTGAATGTAAACAAGT 2739

QY 249 LeuGlnIlyThrMetAsp-----ThrValAsnThrValMetIleAlaValSerVal 265
|          |||          |||          |||          |||
Db 2740 GCACACAGTACATTAAGTACAGTATTCGCGTAATATCGGTAATGTATACAGCAAAATATCT 2799

QY 266 AlaIleThrValIleSerIleValAlaIleIlePheThrLySglnAlaGlyLeuAlaGly 285
|          |||          |||          |||          |||
Db 2800 GCGCATTTAACTGTGTAAGATGCCGCAAAATATGATCGACAGCA----- 2844

QY 286 LeuAlaIleGlyAlaIleValaIleGlyAlaIleAlaIleGlyAlaIleGlyAlaIleAlaIle 305
|          |||          |||          |||          |||
Db 2845 -----GGAGCGCGGACC 2856

QY 306 AlaThrThrValAlaThrGlnIleThrValGlnAlaValAlaValGlnAlaValLySglnAla 325
|          |||          |||          |||          |||
Db 2857 CTAATCGACATCGGGCAATTAACCACTTAAGGCT-----AGTTCAGC 2901

QY 326 ValIleThrAlaValaIlyGlnAlaIleThrAlaIleIleYsaIleAlaValLySleuGly 345
|          |||          |||          |||          |||
Db 2902 ATTAATTCAGCTAATTAACCAAGTAAACCTTCAAGTACAGTGTACATTTGGGGGAAT 2961

QY 346 lIleYsaIlePheIleIlyThrLeuVallySaIleAlaIleAlaYsaAlaIleSerLySgIlyIle 365
|          |||          |||          |||          |||
Db 2962 ATCAATGCTGTATATGTAACTGAATATCTACAGGCGGTCTAATCTACCGTAAGGTTCA 3021

QY 366 SerLySAlaPheAlaLySgIlyThrGlnMetIle-----AlaIlyAsnPheProLySleuSerLyVal 386
|          |||          |||          |||          |||
Db 3022 AGCATTAACGCAACAGCGCACCTGTGTATTAACGCAAAAGAGCGCTGAGCTTAATGT 3081

QY 377 -----AlaIlyAsnPheProLySleuSerLyVal 386
|          |||          |||          |||          |||
Db 3082 GAGGCATCAGTAACCATACAGTACAGTGAATGCAACCAACGCAAAATGCTCCGCGCAGCTA 3141

QY 387 lIleSer-----SerLeuThrSerLySThrValIlyThrVal---GlyVal 399
|          |||          |||          |||          |||
Db 3142 ATCGGCAACCTCAAGCAGAGTGAACATCACTGGGGATTTAATACACATTAATGAATTA 3201

```


Db	2473	CAGGCCCGTTGGAAACGCCACACCGTGG-----ATC	2502
Qy	411	GLYIleMetGlnMetGlnIleuSerGluMetGlnGlnAsnValAlaGlnPheGlnLysGlu	430
Db	2503	GGGGTCGACGACGACAGAGATCTCCAGGTGGGGGCTTCGACCTGGCCGCATTACCGGCAC	2562
Qy	431	ValGlyLysIleuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPheThrGlnGln	450
Db	2563	GGCGACACCTGAATATCCGTTCGACGTTTCGTGGACCCGCTTCGTCAGGGCGACGGGGTG	2622
Qy	451	AlaSerLys-----IleAlaSerLysGln	458
Db	2623	GCCGGCCGACGGCGGTGCTTCGACGTCGCCGCCGGCGGCGCATGGATGCCACGCCGCAAG	2682
Qy	459	ThrGlyGluSerAsnGluMetThr-----	466
Db	2683	--GGCGAGCGCGCGCGCGTCGACGTTTCATACGCCCTGTGGCGCGCAAGAGAAGACGAG	2739
Qy	467	GlnLysAlaThrLysLysIleuGlnIleIleLysAlaThrAlaIleSerGly	485
Db	2740	CGCGGGCGCGAAGAACGGGCAAGACGCAATTCACCAATTCGTGCAGATCTGTGGCG	2796

```

RESULT 42
US-08-669-785-1
: Sequence 1, Application US/08669785
: Patent No. 6309648
:
GENERAL INFORMATION:
:
APPLICANT: Betsou, Potini
:
APPLICANT: Sebou, Peter
:
APPLICANT: Guiso, Nicole
:
TITLE OF INVENTION: Protective Epitopes Of Adeny1
:
TITLE OF INVENTION: Cyclase-Haemolysin(AC-Hly), Their Application To
:
TITLE OF INVENTION: The Treatment Or To The Prevention Of Bordetella
:
Infections

```

CORRESPONDENCE ADDRESS:
 ADDRESSEE: Finegan, Henderson, Farabow, Garrett & Dunner
 STREET: 1300 I Street, N.W., Suite 700
 CITY: Washington
 STATE: D.C.
 COUNTRY: USA
 ZIP: 20005-3315
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/669,785
 FILING DATE: 27-JUN-1996
 CLASSIFICATION: 424
 ATTORNEY/AGENT INFORMATION:
 NAME: Meyers, Kenneth J.
 REGISTRATION NUMBER: 25,146
 REFERENCE/DOCKET NUMBER: 02356, 0072-00000
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (202) 408-4400
 TELEFAX: (202) 408-4000
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 6441 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 979..6096
 OTHER INFORMATION: /note="AMINO ACID SEQUENCE
 OTHER INFORMATION: CORRESPONDING TO THE NUCLEOTIDE SEQUENCE OF THE GENE
 OTHER INFORMATION: CODING FOR THE B. Pertussis AC-Hly"
 US-08-669-785-1

Alignment Scores:	
Pred. No.:	0.000204
Score:	147.50
Percent Similarity:	32.47%
Best Local Similarity:	19.69%
Query Match:	6.23%
DB:	4
Length:	6444
Matches:	114
Conservative:	74
Mismatches:	210
Indels:	181
Gaps:	22

US-09-889-314-2 (1-496) x US-08-669-785-1 (1-6441)

Qy	12	ProAspAsnGlnIlysAsnIleMetSerGlnValLeuThrSerThrProGln-----	28
Db	2083	CCGAGCGGACGTTCCGAAGTTCTCGCCGAGATCTGGAACACGGCCGCCGCTCACCCCGGA	2142
Qy	29	-----GlyValProGlnGlnAspLysLeuSerGlnAsnGluThr	41
Db	2143	TTGCGGCGCCCTGCTGGCGGACGAGTGAACGCCAGAGAT-----TCCGGCTATGACACG	2196
Qy	42	LysGlnIleGlnGlnThrArgGlnGlnIlyLysAsnThrGluMetGluSerAspAlaThrIle	61
Db	2197	CTTGATAGGGGTGGGATCCGATCGTTCTCTGTTGGGAGAGTG--TCCGACATGGCCGC	2253
Qy	62	AlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGln	81
Db	2254	GTCGAAGCGCGCAACTCGAATAAGACC-----CGG	2283
Qy	82	GlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGly	101
Db	2284	CAAGTCTTGACGCCGGGGCGCGGCAGACGAT-----GCCAGCCGGC	2328
Qy	102	ValSerGlyAlaAlaAla-----Thr	108
Db	2329	GTCAGCGGTCCGTCCGGCGCACTGGGGGACGGGGCGCTGCAGGGCGCCAGCGCGTGGCG	2388
Qy	109	ThrAlaSerAsnThrAlaThrLysIleAlaMetClnThrSerIleGluAlaSerLys	128
Db	2389	GCGCGCAGCGCGCTGTTCACTGCATTGCCCTTATGCAATTCGGCGCGCGCGTTC	2448
Qy	129	SerMetClnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla-----GlnMetLys	146
Db	2449	ACC-----AACACGCCCGCAGAAAGCGGCTCTGTTCGGCGGCGCGTTCGGCTTGGCG	2502
Qy	147	GluValClnLysAlaValValAlaAlaAlaLeuSerGlyLysSerGlySerAlaLysLeu	166
Db	2503	GAGGCCAGCAGCGCGGTGGCGGAACCGTAGCGGTTTTCGCGGCGTCTTCGCGCTCG	2562
Qy	167	GluThrProGluLeuProLysProGlyValThrProArgSerGluValIleGluIleGly	186
Db	2563	GCCGGC-----GGT	2571
Qy	187	LeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGluAlaThrLysSerAlaLeuSerAsn	206
Db	2572	TTGCGCGTGGCTGGCGGCGCGGATGGCCTGGAGAGCGGCATCGCCGCGCGCTTGGC---	2628
Qy	207	ThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGluLysGlnAla	226
Db	2629	-----GCCGGC	2634
Qy	227	IleLysIleAspLysGluArgGluGluThrTyGlnGluMetLysAlaAlaGluGlnLysSer	246
Db	2635	ATGTCGTTGACCGATGACGGCGCGCGGACAGAAAGCCCGCGCGCGCCGCAATCGCG	2694
Qy	247	LysAspLeuGlnGluThrMetAspThrValAsnThrValMetIleAlaValSerValAla	266
Db	2695	CTGCAGTTGACAGGT-----GGAACGCTCAGCGTGGCTTCTTCATCGCG	2739
Qy	267	IleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeu	286
Db	2740	TTGGCGCTGGCGCG-----GCCGCGGCGGTACCAAGCGCTTGAGAGTGGCGCGGCG	2793
Qy	287	AlaAlaGlyAlaAlaValGlyAla-----	294
Db	2794	TTGGCGCGGGCGGTGGCGGCGCATTTGGCGCGGCGCTCACTCCATGAGAGATTTACGCG	2853

QY 216 GlnThrAsnLysLeuGlnIleLysIleAspLysLsuArgIleGlu 235
Db 411 GCAGT----- 416
QY 236 TyrGlnIleMetLysAlaIleGlnLysSerLysAspLeuGlyThrMetAspThr 255
Db 417 -----ATCTCAGCGCTGTAATCGGCATCAAGAGTCCCTGCTAGTCAGATCAAG 467
QY 256 ValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaIle 275
Db 468 AGTACGTACGCTCAGCAGACATCAAGCTTCT-----GAATCG 506
QY 276 IlePheThrCysGlyAlaGlyLeuAlaIleGlyAlaIleValAlaIleGlyAlaIle 295
Db 507 GCATCAACCAAGTGGCTCAGCTCAGCAGCAGCAGCAGCTTCAAGCAGTCAAGT 566
QY 296 AlaIleGlyGlyAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIleAlaIle 315
Db 567 GCTTCAAGCTCAGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 617
QY 316 GlnAlaValAlaGlnAlaValAlaValAlaValAlaValAlaValAlaValAlaVal 335
Db 618 TCGGCATCAACCAAGTGGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 677
QY 336 AlaAla 337
Db 678 AGTCA 683
RESULT 41
5183745-1
; Patent No. 5183745
; APPLICANT: DANCHIN, ANTOINE;GLASER, PHILIPPE;KRIN, EVELYN;
; BARZU, OCTAVIEN;LADAMP, DANIEL;ULMAN, AGNES
; TITLE OF INVENTION: ADENYL CYCLASE DERIVATIVES AND THEIR
; BIOLOGICAL USES
; NUMBER OF SEQUENCES: 13
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/426,541
; FILING DATE: 25-OCT-1989
; SEQ ID NO:1
; LENGTH: 4649
5183745-1
Alignment Scores:
Pred. No.: 0.00012 Length: 4649
Score: 147.50 Matches: 114
Percent Similarity: 32.47% Conservative: 74
Best Local Similarity: 19.69% Mismatches: 210
Query Match: 6.23% Indels: 181
Gaps: 22
US-09-889-314-2 (1-496) x 5183745-1 (1-4649)
QY 12 ProAspAsnGlnLysAsnIleMetSerGlnValLeuThrSerThrProGln----- 28
Db 1288 CCGAGCGAGCTTTCGAAGTTCTCCGCGGATGTAAGAAAGCGGCGGTCAACCCGGA 1347
QY 29 -----GlyValProGlnGlnAspLysLeuSerGlyAsnGluThr 41
Db 1348 TTGCGCGCGCGCTCGCTGCGCGCGAGTGAGACGCCAGAT-----TCCGGCTATGACAGC 1401
QY 42 LysGlnIleGlnIleThrArgIleGlnIleLysAsnThrGlnIleGlnIleGlnIle 61
Db 1402 CTTGATGGGTGGATCGGATCGGATCTTCGTTGGCGGAGTG-----TCCGACATGCGCGC 1458
QY 62 AlaGlyAlaSerGlyLysAspLysThrSerSerThrLysThrGlnThrAlaProGln 81
Db 1459 GTGGAAGCGCGGCAACTGGAATGAC-----CGG 1488
QY 82 GlnGlyValAlaIleAlaIleLysGlnLysSerGlnLysAlaGlyAlaAspThrGly 101
Db 1489 CAAGCTTTCGACGCGCGCGCGCGCGAGACGAT-----GCCAGCGCGGCG 1533

QY 102 ValSerGlyAlaAlaAla-----Thr 108
Db 1534 GTGAGCGGTGCGCGCGCAGCTGGGGCAGCGGCGCTGACGGGCCCCAGCGGTTGGC 1593
QY 109 ThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlnIleAlaSerLys 128
Db 1594 GCGGCGAGCGGCTGCTTCATGCTATGCTGATGACGAATTCGCGCGCGGCTTCC 1653
QY 129 SerMetGlnSerThrLeuGlnLysLeuGlnSerLeuSerLysAla-----GlnMetLys 146
Db 1654 ACC-----AACCGCGCGCAGGAAGCGCGCTGTGCGCGCGCTTCCGCTTGGC 1707
QY 147 GluValGlnAlaValAlaValAlaValAlaValSerGlyLysSerGlySerAlaLysLeu 166
Db 1708 GAGGCGAGCGCGCGCTGCGCGGAAACCTGAGCGGTGTTTCCGCGGCTTCCGCGCTGG 1767
QY 167 GlnThrProGlnLeuProLysProGlyValThrProArgSerGlnValIleGlnIleGly 186
Db 1768 GCGCGC-----GCT 1776
QY 187 LeuAlaLeuAlaLysAlaIleGlnThrLeuGlnIleAlaThrLysSerAlaLeuSerAsn 206
Db 1777 TTGCGGCTGCGCGCGCGCGCGATGCGCTGAGAGCGCGGCGATCCCGCGCGCTTGGC 1833
QY 207 TyrAlaSerThrGlnAlaGlnAlaAspLynThrAsnLysLeuGlnLysGlnAla 226
Db 1834 -----GCCGCG 1839
QY 227 IleLysIleAspLysGlnArgIleGlnIleMetLysAlaAlaGlnLysSer 246
Db 1840 ATGCTGTTGACCGATGACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1899
QY 247 LysAspLeuGlnIleThrMetAspThrLysAsnThrValMetIleAlaValSerValAla 266
Db 1900 CTGACATTGACAGT-----GGAACCGTGCAGCTGCTTCTTCAATCGCG 1944
QY 267 IleThrValIleSerIleValAlaIlePheThrCysGlyAlaGlyLeuAlaIle 286
Db 1945 TTGCGCGTGGCGCG-----GCCGCGCGCTGACCAAGCGCTTGGAGTGGCGCGCGCG 1998
QY 287 AlaIleGlyAlaAlaValAlaIle-----GAAACCGTGCAGCTGCTTCTTCAATCGCG 294
Db 1999 TCGGCGGCGCGCGCTGCG 2058
QY 294 ----- 294
Db 2059 CTGCTGACAGATGACACTATGCGATGACGTGACAGCTGGCGCAGATGACGCGCA 2118
QY 295 -----AlaAlaAla 297
Db 2119 TACGCTTACGAGGCGGAGCGCTTGTGCGCCAGCTGATGCGCAGACAGAGCGCGCGAG 2178
QY 298 GlyGlyAlaIleGlyAlaAlaAlaAla-----ThrThrValAlaThrGlnIleThrValGln 316
Db 2179 GCGCGCGTCCGCGCGCTCCGCTGAGCAGCGGTGGCGCGCGCGCGCGCGCGCGCGCG 2238
QY 317 AlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAla 336
Db 2239 GCGCGCGCGCGCGGTGA-----CGGCGCGCGGTGGCGGTGACCTTCTTGGTACCGCG 2295
QY 337 AlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAla 356
Db 2296 GCTCTCAACGCACTCTGCGCGCGGTGACGAGCCATCATGAAAGCTGCAACGAT 2355
QY 357 IleAlaLysAlaIleSerLys-----GlyIleSerLysValPheAlaLysGlyThrGln 374
Db 2356 TACGCTTCGAAGTGCAGAGGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2415
QY 375 -----MetIleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeu 390
Db 2416 GCGCGTCAAGCAACTGCGC-----AATTCGAGCGCGCTTACGGAATGCTGGCGCGACTG 2472
QY 391 ThrSerLysThrValThrValGlyValAlaValAlaIleProAlaLeuGlnLys 410

[illegible]

```

RESULT 40
US-08-961-527-377
: Sequence 377, Application US/08961527
: Patent No. 6420135
: GENERAL INFORMATION:
: APPLICANT: Charles Kunsch
: TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
: NUMBER OF SEQUENCES: 391
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue

```

```

CITY: Rockville
STATE: Maryland
COUNTY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MD storage
COMPUSER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEO ID NO: 377 :
SEQUENCE CHARACTERISTICS:
LENGTH: 690 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-377

Alignment Scores:
Pred. NO.: 5.44e-06 Length: 690
Score: 147.50 Matches: 62
Percent Similarity: 37.23% Conservative: 43
Best local Similarity: 21.99% Mismatches: 112
Query Match: 6.23% Indels: 65
DB: 4 Gaps: 5

US-09-889-314-2 (1-496) x US-08-961-527-377 (1-690)
Oy 56 GluseRaspAlatHrIleAalaglylasergilylsasplysrThrseserThrhlrllys 75
    ||||| .....||| | ||||| ||||| ..:::.....
Db 33 GAATGCGCATCAAGACAGTGTCTGGCTTCACATCAACAAGTGCGTTACGCCAAC 92
Oy 76 ThrGlutrrAlaProGLngInglNylValAlalaAgLyLysgLuSerSerglUsrglnlys 95
    ::::: ||||| ....:..||| ||||| .....
Db 93 ACATCAGCTTCTGAATCGATCGCATCAACCAGTCGCCTTCACAGCTCAACCAATGGCTG 152
Oy 96 AlactylaAsPrhgrylvalSerglyVAlAlaAlaAlaThlrThralasernThrlaThlr 115
    ||| ||| ||| ||| ||| :::||::|||::||| ||| ::|||:::
Db 153 GCTTCAGCGCTGCAACAAGTGTCTCCGCTTCACATCAACAAGTGGCTGCCTCACGACG 212
Oy 116 LyslIleAlamEtGLnTrserIIleGLugluAlaserlysSemetgLuserrThrlenglu 135
     ||||||| ..... ||||| ||| .....
Db 213 -----||| ||||| ||||| .....
Oy 136 SerLeugInserLeuSerAlaAlaGLmetLsgLuVaAlguAlaValAlaValAla 155
   ::: ||| |||||::: ||| ||| :::: |||:::
Db 243 ACCAGTGGCTCGCTCAGCAAGCAACAAGTCGCTACAGCAATATCTCAAGCTCT 302
Oy 156 LeuSerglyLysSerSerglySerAlaLysLeuGIUthrProGUleuProLysProGI 175
     ||| ::||| |||||
Db 303 GAATGCGCATCAACGAGTGGCTCGGCTCA----- 332
Oy 176 ValThrProArgSerGIuValIlleGIuIlleLysLeuAlaleuAlaLysAlaIlleGINThr 195
     ||| ||| ||| |||
Db 333 -----GCAACGCGAACTACTCA 350
Oy 196 LeuGIyGuAlaThrLysSerAlaLeuSerAsnTyraLaserrThrgInalaglinalasp 215
     ::::: ||| ||||| ||| |||
Db 351 GGCTCAGCTTGCGCTCAACCAAGTGGCTTCAGACAAGCACAACTGGCTCAGCTCA 410
```

```

APPLICANT: Roop, Dennis R.
APPLICANT: Rothnagel, Joseph A.
APPLICANT: Greenhalgh, David A.
TITLE OF INVENTION: CONSTITUTIVE AND INDUCIBLE EPIDERMAL
TITLE OF INVENTION: VECTOR SYSTEMS
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSER: Fulbright & Jaworski
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77010-3095

COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/03993
FILING DATE: 19930428
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-5405
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5325
TELEFAX: 713/651-5246
TELEX: 762829
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6530 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US93-03993-1

Alignment Scores:
Pred. No.: 0 000185 Length: 6530
Score: 148.00 Matches: 92
Percent Similarity: 33.74% Conservative: 45
Best Local Similarity: 22.66% Mismatches: 165
Query Match: 6.25% Indels: 106
Gaps: 10
DB: 5

US-09-889-314-2 (1-496) x PCT-US93-03993-1 (1-6530)
QY 37 sergIyasnclutfrlysglnllieglnglntfhrarglnelylsasnthrlumetlu 56
Db 2548 AAAGGAGTCCACACACTTAGTCGTGCATTAACCTAGGAGTGAGAAGATATGTGGCATGAT 2607
QY 57 SerAspAlaThrIleA----- 62
Db 2608 GCATCGGCCACCCTTACACAGCGCTCTGCTGCTGTGGTCAATGTTGCTCTTCGTC 2667
QY 63 -----GLYAlaserlylysaspIysrserThrrlrlystrgluthrala 79
Db 2668 TTCCTTCAGGGTTCCTCTTCCTTA-AAACAAGATGTCACCGAAMAAAGCAGCCACT 2722
QY 80 ProIngInsllyValAlaIlaaglylsgIuseSer-GIuseRglInylsAlaIylalAs 99
Db 2727 CCGTCGCCCTCCTCGGTGGTTGTGAAGAACCTCTGTTGGAGGAGAGAGCGGCGGCTAT 2786
QY 99 pthrIglyvalserGIyAlaAlaIathrrThrIalasernthrAlathrrysIleAlawe 119
Db 2787 TATAGCGGTGGCGGCTCTGCGTGGAGGCGGCTAT----- 282
QY 119 tGlntfhrserlleaglInulaIaserlysermetgluseThrrleugluseIleuglne 139

```


Db	2727	CCGAGCCCTCTCTGTGGGTTGTGGAAGACCTCTGCTCGAGGAGAGAGGCGCGCGCTAT	2786
Oy	99	pThrGlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaIle	119
Db	2787	TATAGCGGTGGCGGCTCTGCTCGGAGGCGGCTCAT	28233
Oy	119	tGlnThrSerIleGluGlnLysLysSerMetGluSerThrLeuGluSerLeuGlnSe	139
Db	2824	-----CTGGAGGAGCGTCTAGC	2840
Oy	139	rLeuSerAlaAlaGlnMetLysGluValGlnAlaValAlaValAlaAlaLeuSerGly	159
Db	2841	TCGTGAGGCGGAGCGCGGTGTCTTCATTGAGAGCGTCTCCAGCTGGCGCGGTGAGAGCGCGC	2900
Oy	159	sSerSerGlySerAlaLysLeuGluThrProGlu-----LeuProLysProGlu	175
Db	2901	TCCGGTGGGGCGCTCA-----AGTACTCCGGAGCGCGGTGCGCTACAGTCGGCGCGC	2954
Oy	175	yValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThr	195
Db	2955	GGCTACTCCGAGGCGGTGTG-----GCTCTAGCTGGCGCGGTGCTAC	2999
Oy	195	rLeuGlyGluAlaThrLysSerSerAlaLeuSerAsnTyrrAlaSerThrGlnAlaGlnAlaAs	215
Db	3000	TCTGGGGCGCGCGCGCGCTCCACT---CGAGAGTGGCTACTCCGAGAGCGCGCGCGC--	3054
Oy	215	pGlnThrAsnLysLeuGlyLeuGluLysGlnAlaIleLysIleAspLysGluArgGlu	235
Db	3054	-----	3054
Oy	235	uTyrlGlnGluMetLysAlaAlaGluGlnLysSerLysAspLeuGluGlyThrMetAspTh	255
Db	3055	-----GCTCCAGCTCGCGCGCGGCGAGCTACTCCGGGGGTGCTCCAGC	3098
Oy	255	rValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAl	275
Db	3099	TGTGGAGCGCGTGGCGCGCTGTGGTGGGGCGCGTCAACTCCGAGAGTGTGGCGCGCGC	3158
Oy	275	aIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAla-----	291
Db	3159	GGCTCT-AGCTGCGCGCGCGCTCTCCGGGGCGCGCGCGCTCCAGCTCGCGGAGG	3217
Oy	292	-----ValGlyAlaAlaAlaAlaGlyGlyAlaAlaGly-A	303
Db	3218	CGGATCAGGAGCGCGCGCTACTGCGGAGGCTCTCGAGGCGGCGAGCTCCGGTGG	3277
Oy	303	lAlaAlaAlaAlaThrThrValAlaAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValL	323
Db	3278	CTCGCGCGCGCGCTCCGAGGCGCGCAAGTACTGTGTGGCGCGGTGGCTCCAGCTCGCGG	3337
Oy	323	ysGlnAlaValIleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaAlaValL	343
Db	3338	AGGCGCGTATTCGCGCGCGGTGAAGACAGCGCGCTCACTGATGGCGGCGCTACTC	3397
Oy	343	ysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerL	363
Db	3398	AGGTGGCGGTGATCA-----GCTGGCGCGCGCGG	3430
Oy	363	ysGlyIleSerLysValAlaPheLysGlyThrGlnMetIleAlaLysAspHeProLysL	383
Db	3431	CGGCTATTCCGTGGCGCGCGCACAGACTCGGAGGTGT-----CCTCCGG	3478
Oy	383	euserLysValIleSerSerLeuThrSerLysTrpValThrValGlyValGlyValVal	403
Db	3479	TGGCGGCGCGCGAGATCGTCCACAAGATTCAGTCCAGACTACGAGAGGCGGTTCTAG	3538
Oy	403	alaAlaAlaProAla	407
Db	3539	CGGTGGCTCCAGT	3552
RESULT	38		
US-08-458-240-1			
Sequence 1, Application US/08458240			

```

Patent No. 6143727
GENERAL INFORMATION:
APPLICANT: Roop, Dennis R.
APPLICANT: Rochnagel, Joseph A.
APPLICANT: Greenhalgh, David A.
TITLE OF INVENTION: SPECIFIC EXPRESSION VECTORS
TITLE OF INVENTION: AND METHODS OF USE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: LYON & LYON
CITY: Los Angeles
STREET: 611 West Sixth Street
STATE: California
COUNTRY: U.S.A.
ZIP: 90017
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: IBM MS-DOS (Version 5.0)
SOFTWARE: Wordperfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,240
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/146,930
FILING DATE:
APPLICATION NUMBER: 07/876,286
FILING DATE: April 30, 1992
APPLICATION NUMBER: NO. 6143727 yet assigned (204/132)
FILING DATE: October 29, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 204/152
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6530 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-458-240-1

Alignment Scores:
Pred. No.:          0.000185      Length:    6530
Score:              148.00        Matches:   92
Percent Similarity: 33.74%       Conservative: 45
Best local Similarity: 22.66%     Mismatches: 165
Query Match:         6.25%        Indels:    106
DB:                  3           Gaps:       10

US-09-889-314-2 (1-496) x US-08-458-240-1 (1-6530)
Oy      37 serglynsgluThrLysGlnIleGngInlThrArgGlngLyAsnThrcJumetglu 56
::||| |::||| |::||| |::||| |::||| |::||| |::||| |::||| |::|||
Db      2548 AATGGAGTCGCACGACCCTAATGTGCAATAAGACTAGSAGTCCACAAGAATATTGGCATGGAT 2607
                    |||||||
Oy      57 SeraspAlaThrIleala----- 62
::: | |||||
Db      2608 GCATCTGCCACCTTCACAGCGCTCCTTGCTGAATGTTAGTTCCTTCGCTC 2667
                    |||||
Oy      63 -----glyAlaserGlyLyasplysthSerSerThrlThlysthgCluthra 79
||| ::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::| |::|
Db      2668 TTCCTTCAGGGGTTCCCTTCCTTA-AACAAAGATGTCACACAAAAGAGCCCACT 2726
ProGlnGlnGlyValaIalaaIagLyLysGluSerSer-GluSerGlnLyAlaGylAlas 99
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2727 CCCTGCCCTCTCTGTGGGTTTGGAAGAAGACCTTGtGGAgAGAGAGcGcGcgccgcctAT 2786

```



```

OY 280 GLYAlaGlyLeuAlaGlyLeuAlaGlyAlaAlaValAlaAlaAlaGlyGly 299
DB 1391 AGCACAATATTAGTGGC---TCAACAAGTACAGAGGTTTCAGACTCAACAAGTACATCG 1335
OY 300 AlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 319
DB 1334 ACATCAGACAGTACATTCAGTCAACA-----AGTGAAGGCGAGTCCGACAGTGAAGT 1281
OY 320 GlnAlaValLysGlnAlaValLLeuAlaValAlaAlaAlaAlaAlaAlaAlaAlaAla 339
DB 1280 ACGTCATCAAGTAAATCAAGTACAGCAGTTCAGATTCAACAAGCGCTCAACAGTCA 1221
OY 340 AlaAlaValLysSerGlyLLeuLysAlaPheLeuLysThLeuValLysAlaAlaAlaLys 359
DB 1220 GAAAGTGGCTCAACATCAACAGAGTGTG-----AGTACTCGAATATAGTCAAGTACG 1170
OY 360 AlaLeuSerLysGlyLLeuSerLysValPheAlaLysGlyThrGlnMetIleAla----- 377
DB 1169 TCATTAAAGTAAATCAACAAGTACAGTCTTCAGATTCAAGAGTATGTCAGATCAGAC 1110
OY 378 -----LysAsnPheProLysLeuSerLysValLLeuSer 389
DB 1109 AGTGGTCACATCAACAAGTACAGAGCGAGTCTCGACAGTCAAGTACATCGTTAAGTGA 1050
OY 390 LeuThrSerLysTrpValThrValAlaGlyVal-----ValValAla 404
DB 1049 TCGCAAAAGTACAGAGGTTTCAGAAATCGACAACTACATCAACAAGTGTGAGTGGC 990
OY 405 AlaProAlaLeuGlyLysGlyLLeuMetGlnMetGlnLeuSerGlnMetGlnLysAla 424
DB 989 TCAAAACAGTACAAAGTACGATTAAAGTACGTCAGAAAGTACAGTCTTCACAGATCAAG 930
OY 425 AlaGlnPheGlnLysGlyValGlyLysLeuGlnAlaAla---AlaAspMetLeuSerMet 443
DB 939 ACCACATCGACATCAAGTACGATCAACGTCACAAAGGAGGACGACGTCGAGCACTGCA 870
OY 444 PheThrGlnPheThrGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyLysLeu 462
DB 869 AGTACATCTTAAAGTACGATCAACAAGTACAAAGTTCAGACTCAACGAGTACATCA 813

RESULT 36
US-08-669-785-3
: Sequence 3, Application US/08669785
: Patent No. 6309648
: GENERAL INFORMATION:
: APPLICANT: Betsou, Fotini
: APPLICANT: Sebo, Peter
: TITLE OF INVENTION: Protective Epitopes Of Adenyl
: TITLE OF INVENTION: Cyclase-Haemolysin(AC-Hly), Their Application To
: TITLE OF INVENTION: The Treatment Or To The Prevention Of Bordetella Infections.
: NUMBER OF SEQUENCES: 4
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Finegan, Henderson, Farabow, Garrett & Dunner
: STREET: 1300 I Street, N.W., Suite 700
: CITY: Washington
: STATE: D.C.
: COUNTRY: USA
: ZIP: 20005-3315
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/669,785
: FILING DATE: 27-JUN-1996
: CLASSIFICATION: 424
: ATTORNEY/AGENT INFORMATION:
: NAME: Meyers, Kenneth J.
: REGISTRATION NUMBER: 25,146
: REFERENCE/DOCKET NUMBER: 02356, 0072-00000

```

```

: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (202) 408-4000
: TELEFAX: (202) 408-4400
: INFORMATION FOR SEQ ID NO: 3:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 5118 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: DNA (genomic)
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 1..5115
: OTHER INFORMATION:
: OTHER INFORMATION: CORRESPONDING TO THE NUCLEOTIDE SEQUENCE OF THE GENE
: OTHER INFORMATION: CODING FOR THE B. Bronchiseptica AC-Hly"
US-08-669-785-3

Alignment Scores:
Pred. No.: 7,83e-05 Length: 5118
Score: 150.00 Matches: 116
Percent Similarity: 33.04% Conservative: 74
Best Local Similarity: 20.17% Mismatches: 211
Query Match: 6.33% Indels: 174
DB: 4 Gaps: 22

US-09-889-314-2 (1-496) x US-08-669-785-3 (1-5118)
OY 8 SerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeuThrSerThrPro 27
DB 1117 TCGAAGTCCCTCCCGGATGCTACTGGAACGGTCCGCGCTCACCCGAGTTGGCGGCGG 1176
OY 28 Gln-----GlyValProGlnLysAspLysLeuSerGlyAsnGlnThrLysGlnIleGln 45
DB 1177 TCGCTGGCGGCGAGTGAACGACGAGAT-----TCCGCTATGACAGCTTGATGGGCTG 1230
OY 46 GlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGlyAlaSer 65
DB 1231 GGATCGCGATCGTTCCTGTCGTCGCGGCGAGTGC---TCCGATATGGCCCGCTGGAACGGCG 1287
OY 66 GlyLysAspLysThrSerSerThrLysThrLysAlaProGlnGlnIleValAla 85
DB 1288 GAACGCGAATATGACC-----CGCAAGCTTTCAC 1317
OY 86 AlaGlyLysGlnSerSerLysLeuLysAlaGlyAlaAspThrGlyAlaSerGlyAla 105
DB 1318 GCCGGCGCGCGCAGACGAT-----GCCGACCGGCGCTGAGCGGTGGC 1362
OY 106 AlaAla-----ThrAlaSerAsn 112
DB 1363 TCGGCGCACTGGGGCGAGCGGCGCTGCAAGGCGCCAGCGCGGTGGCGGCGCAGCGG 1422
OY 113 ThrAlaThrLysIleAlaMetGlnThrSerIleGlnAlaSerLysSerMetGlnSer 132
DB 1423 CTGGTTCAATGATCCCTGATGACGCAATTCGGCGGCGGCTTCACAC-----AAC 1476
OY 133 ThrLeuGlnSerLeuLysLeuSerAlaAla-----GlnMetLysLeuValGlnAla 150
DB 1477 ACGCGCAGGAAGCGCGCTCTGTCGGCGCGCTTCGCTTGGCGGCGACACACAC 1536
OY 151 ValValAlaAlaAlaLeuSerGlyLysSerSerGlySerAlaLysLeuGlnThrProGlu 170
DB 1537 GCCGTGGCGGAACCGTGAAGCGGTTTTCGCCGGGTCTTCGCCGTGGCGGCGG----- 1590
OY 171 LeuProLysProGlyValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAla 190
DB 1591 -----GGTTTCGGCGTGGCT 1605
OY 191 LysAlaIleGlnThrLeuGlnLysAlaThrLysSerAlaLeuSerAsnThrAlaSerThr 210
DB 1606 GCGCGCGCGATGGCGTGGGA----- 1626
OY 211 GlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLysGlnValIleLysIleAsp 230

```

```

Db 3091 ACATCCGACGATCATCAATGTCACGACCGAGAGTGAATCAACAGT----- 3138
      |||::: |||::: |||
      |||::: |||::: |||
Qy 224 LysGlnAlaIleLysIleAspLysGluArgGluGluThrGlnGluMetLysAlaIleGlu 243
      |||::: |||::: |||
Db 3139 -----AAAAGTACGTCATTAAAGTCGAA 3159
      |||
Qy 244 GlnLysSerLysAspLeuGluGly-----ThreMetAspThrValAsnThr 258
      ||| ||| |||
      ||| |||::: |||::: |||
Db 3160 TCGACAGTACGATCTTTTCAGGCTCAACGAGTGCCTCGACATCCGACGATGATCAACG 3219
      ||| ||| |||
      ||| |||::: |||
Qy 259 ValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaIleAlaIlePheThr 278
      ||| ||| |||
      ||| |||::: |||
Db 3220 TCAACGAGTACGAGTACGATCAGACAGTACAGTACAGTACATTAAGTGAATGACAAAGTACA 3279
      ||| ||| |||
      ||| |||::: |||
Qy 279 CysGlyAlaGlyLeuAlaGlyLeuAlaIleGlyAlaIleValGlyAlaIleAlaIleGly 298
      |||::: |||::: |||
      |||::: |||::: |||
Db 3280 -----AGTCTTTCAGGCTCAACGAGTGC 3303
      |||
Qy 299 GlyAlaAlaGlyAlaAlaAlaAlaIleThrThrValAlaIleThrGlnIleThrValGlnAlaVal 318
      |||::: |||::: |||
      |||::: |||::: |||
Db 3304 TCGACATCCGACAGTCAATCAACGTCACAGCAGAGTACATCAACAGTGAAGTACA 3363
      |||::: |||::: |||
      |||::: |||::: |||
Qy 319 ValGlnAlaValLysGlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAlaIle 338
      |||::: |||::: |||
      |||::: |||::: |||
Db 3364 -----TCGTTGAGTATCATTAAGCAGCAGCGTTTCAGATTCAGACAGCGCTCAACG 3417
      |||::: |||::: |||
      |||::: |||::: |||
Qy 339 ---LysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIle 357
      |||::: |||::: |||
      |||::: |||::: |||
Db 3418 TCGAAAGTGCATCCACATCAACGAGTACAGTCAACAGTCAACAGTCAACATCGCTTA 3477
      |||::: |||::: |||
      |||::: |||::: |||
Qy 356 AlalysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAla 377
      |||::: |||::: |||
      |||::: |||::: |||
Db 3478 AGTGCATCTTAAAGTACAAAGCATTTTCAGATTCACAGATACGTCAGATCCGACAGTCA 3537
      |||::: |||::: |||
      |||::: |||::: |||
Qy 378 LysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysThrValThrVal 397
      |||::: |||::: |||
      |||::: |||::: |||
Db 3538 TCACAGCTCAACGAGTACGATGAGTACAGCAGTACACAGTACATCACTCA----- 3582
      |||::: |||::: |||
      |||::: |||::: |||
Qy 398 GlyValGlyValValValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGln 417
      |||::: |||::: |||
      |||::: |||::: |||
Db 3583 -----TTA 3585
      |||
Qy 418 SerGlnMetGlnGlnAsnValAlaGlnPheGlnLysGluValGlyLysLeuGlnAlaIle 437
      |||::: |||::: |||
      |||::: |||::: |||
Db 3586 AGTGAATCGACACAGCACA-----AGTCTT 3609
      |||::: |||::: |||
      |||::: |||::: |||
Qy 438 AlaAspMetIleSerMetPheThrGlnPheTrpGlnAlaSerLysIleAlaSerLys 457
      |||::: |||::: |||
      |||::: |||::: |||
Db 3610 TCGGACTCAACAGTACGTCACAG-----TCAGAAAGTGCATCAACAG 3651
      |||::: |||::: |||
      |||::: |||::: |||
Qy 458 GlnThrGlyLysSerAsnGlnMetThrGlnLysAlaThrLysLeu 472
      ||| ||| |||
      ||| |||::: |||
Db 3652 TCAACGAGTGAAGT---GACTCAACAGTGAAGTACATCACTTA 3693
      ||| ||| |||
      ||| |||::: |||

```

```

; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1685

Alignment Scores:
Pred. No.: 3,87e-05 Length: 3561
Score: 150.50 Matches: 89
Percent Similarity: 37.47% Conservative: 98
Best Local Similarity: 17.84% Mismatches: 241
Query Match: 6.35% Indels: 71
DB: 4 gaps: 14

US-09-889-314-2 (1-496) x US-09-134-001C-1685 (1-3561)

Qy 2 ThrAsnMetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGln 21
      |||::: ||| ||| |||::: |||::: |||
      |||::: |||::: |||
Db 2210 ACGAGTACGTCGACATCCGACAGTCAACGTCACAGTCAACGAGTGAAGTACATCAACAGT 2151
      |||::: |||::: |||
      |||::: |||::: |||
Qy 22 ValLeuThrSerThrProGlnGlyValProGlnAspLysLeuSerGlyAsnGluThr 41
      |||::: |||::: |||
      |||::: |||::: |||
Db 2150 GAAAGTACGTCATTAAAGCAGTCGATTA-----AGCACAGCGCTTCAGATTCAAACAGC 2097
      |||::: |||::: |||
      |||::: |||::: |||
Qy 42 LysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGluSerAspAlaThrIle 61
      |||::: |||::: |||
      |||::: |||::: |||
Db 2096 GCGTCACATCCGACAGTACGTCACATCAACGAGTGAAGTCAAGTCAAGTCAAGT 2037
      |||::: |||::: |||
      |||::: |||::: |||
Qy 62 AlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGln 81
      |||::: |||::: |||
      |||::: |||::: |||
Db 2036 ACATCGTTAAAGTGCCTCGACAAAGTACAGTCTTCAGATTCACAGATCAACATCGACATCA 1977
      |||::: |||::: |||
      |||::: |||::: |||
Qy 82 GlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyValAspThrGly 101
      |||::: |||::: |||
      |||::: |||::: |||
Db 1976 GAC-----AGTCAATCAACGTCACACAGTGAAGGAGTCCGAAAGACCAAGTACATCC 1923
      |||::: |||::: |||
      |||::: |||::: |||
Qy 102 ValSerGlyAlaAlaAlaIleThr-----ThrLysSerAsnThrAlaThrLys 116
      |||::: |||::: |||
      |||::: |||::: |||
Db 1922 TTAAGTGCCTCGACAAAGTACAGTCTTCAGATTCACAGATCAACATCGACATCAACAGTCA 1863
      |||::: |||::: |||
      |||::: |||::: |||
Qy 117 IleAlaMetGlnThrSerIleGlnGlnAlaSerLysSerMetGluSerThrLeuGluSer 136
      |||::: |||::: |||
      |||::: |||::: |||
Db 1862 GCATCAACATCAACAGAGTGAAGTGAAGTCAACAGTGCAGACATCAATCGATTAAAGTGCCTCA 1803
      |||::: |||::: |||
      |||::: |||::: |||
Qy 137 Leu-----GlnSerLeuSerAlaAla 143
      ||| ||| |||
      ||| |||::: |||
Db 1802 TTAAGTACAAAGCGTTTACAGCTCAACAGTACATCAACATCGATGATGATCGATCA 1743
      |||::: |||::: |||
      |||::: |||::: |||
Qy 144 GlnMetLysGluValGlnAlaValAlaAlaAlaLeuSerGlyLysSerSerGlySer 163
      |||::: |||::: |||
      |||::: |||::: |||
Db 1742 ACAAGTGAAGCAGCAGCTCGAAAGACAAAGTACATCGTTGAGTGCCTCAACAGTACAGC 1683
      |||::: |||::: |||
      |||::: |||::: |||
Qy 164 AlalysLeuGluThrProGlnLeuProLysProGlyValThrProArgSerGluValIle 183
      ||| ||| |||
      ||| |||::: |||
Db 1682 ATTTCAGACTCAACAGTACATCAACATCAACAGTGCATCAACATCAACAGTGTGAGT 1623
      ||| ||| |||
      ||| |||::: |||
Qy 184 Glu-----IleGlyLeuAlaLeuAlaLysAlaIleGlnThr---LeuGlyGlnAla 199
      ||| ||| |||
      ||| |||::: |||
Db 1622 GAGTCAAAAGTACAGACATCCATTAAGTGAATCATTAAGTACAGACGTTTCAGATCA 1563
      ||| ||| |||
      ||| |||::: |||
Qy 200 ThrLysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLys 219
      ||| ||| |||
      ||| |||::: |||
Db 1562 ACGAGTACATCGACATCAACAGTGAAGTCAACATCAACAGAGTGTGAGTGCAGC--- 1506
      ||| ||| |||
      ||| |||::: |||
Qy 220 LeuGlyLysGluLysGlnAlaIleLysIleAspLysGluArgGluGlyThrGlnGluMet 239
      ||| ||| |||
      ||| |||::: |||
Db 1506 ----- 1506
      |||
Qy 240 LysAlaAlaGluGlnLysSerLysAspLeuGluGlyThrMetAspThrValAsnThrVal 259
      ||| ||| |||
      ||| |||::: |||
Db 1505 ---AGTCAAGTACGTCATCAACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 1449
      ||| ||| |||
      ||| |||::: |||
Qy 260 MetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCys 279
      ||| ||| |||
      ||| |||::: |||
Db 1448 ACATCCACATCAAGTATGATGATCAATCAACAGAGTGAAGTCAACAGTACAGT--- 1392
      ||| ||| |||
      ||| |||::: |||

```

Oy	145	MettysgluValGlnIvalValValAlaAlaLeuSerGlyLysSerSerGlySerAla	164
Db	2311	AGTAAATCGCAACAGCACACCCTTCACGATTCAACAAGACTGCACATCCGACAGTGCG	2252
Oy	165	LysLeugluThrProGluLeuProLysProGlyValThrProArgSerGluValIleGlu	184
Db	2251	TCCACATCCAACGAGTGAG-----AGGAGTCAGACAGCT	2219
Oy	185	IleGlyLeuAlaLeuAlaLysAlaIleGlnThr--LeuGlyGluAlaThrLysSerAla	203
Db	2218	GCAAGTACGTCAATTAAGTAGAATCGACAAGAATCAAAGCGTTTCAGATTCAACAGACGATCGC	2159
Oy	204	LeuSerAsnTrpLysSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlu	223
Db	2158	ACATCCCAACAGTCATCAATATTCAGGACGCCAGAGTGAATCCAACAGT-----	2111
Oy	224	LysGlnAlaIleLysIleAspLysGluArgGluGluTygGlnGluMetLysAlaAlaGlu	243
Db	2110	-----AAAAGTACGTCATTAAGTGAA	2090
Oy	244	GlnLysSerLysAspLeuGly-----ThMetAspThrValAsnThr	258
Db	2089	TCGACAGTACGAGTCTTTCGGCTCAGACAGAGTGCCTCAGCATCCGACAGTGCATCAACG	2030
Oy	289	ValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaIlePheThr	278
Db	2029	TCCAACGATGAGAGTGCAGTCAGACAGTCAAGTACTCTTAAGTGAATCGCAAGTACG	1970
Oy	279	CysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValaGlyAlaAlaAlaGly	298
Db	1969	-----AGTCTTTCAGGCTCAACGAGTCCG	1946
Oy	289	GlyAlaAlaGlyAlaAlaAlaAlaThrThrValAlaIleThrValGlnAlaVal	318
Db	1945	TCGACATCCCGACAGTGCATCAACGTCACACAGGACGAGTCTCAACAAGTGAAGTACA	1886
Oy	319	ValGlnAlaValLysGlnAlaValIleThrAlaValaArgGlnAlaIleThrAlaAlaIle	338
Db	1885	-----TCGTTGAGTGAATATTATTAAGCACACAGGCTTTCAGATTCAACAAGCCGTCACG	1832
Oy	339	--LysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIle	357
Db	1831	TCACAAGATGTCATCCATCAACAGAGTGAAGTCAAAAGTCCAGCACATCGTTA	1772
Oy	358	AlaLysAlaIleSerLysGlyIleSerLysValaPheAlaLysGlyThrGlnMetIleAla	377
Db	1771	AGTGATCTATTAACTATCAACACATTTTCAGATTCAACGAGTACGTGACATCCGACAGTCA	1712
Oy	378	LysAsnPheProLysLeuSerLysValIleSerSerIleThrSerLysTrpValThrVal	397
Db	1711	TCAACGTCAACGAGTGAAGTGAAGTCAAGACAGTACAGTACAGTCA-----	1667
Oy	398	GlyValGlyValValAlaAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeu	417
Db	1666	-----TTA	1664
Oy	418	SerGluMetGlnGlnAsnValaAlaGlnPheGlnLysGluValGlyLysLeuGlnAlaAla	437
Db	1663	AGTGAATCGACACACACA-----AGTCTT	1640
Oy	438	AlaAspMetIleSerMetPheThrGlnPheTrpGlnAlaSerLysIleAlaSerLys	457
Db	1639	TGCGACTCAACAATACTAGTCAACG-----TCAGAAGTGCATCAACG	1598
Oy	458	GlnThrGlyLysSerAsnGluMetThrGlnLysAlaThrLysLeu	472
Db	1597	TCAACGACTGAGACT---GACTCAACAAGTGAAGTACGTCATTTA	1556
RESULF 34			
US-09-134-001C-1626			
Sequence 1626, Application US/09134001C			
Patent No. 6380370			

```

: GENERAL INFORMATION:
: APPLICANT: Lynn Doucette-Stamm et al
: TITLE OF INVENTION: NOCETIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
: TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: GTC-007
: CURRENT APPLICATION NUMBER: US/09/134,001C
: PRIOR FILING DATE: 1998-08-13
: PRIOR APPLICATION NUMBER: US 60/064,964
: PRIOR FILING DATE: 1997-11-08
: PRIOR APPLICATION NUMBER: US 60/055,779
: PRIOR FILING DATE: 1997-08-14
: NUMBER OF SEQ ID NOS: 5674
: SEQ ID NO 1626
: LENGTH: 6414
: TYPE: DNA
: ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1626

Alignment Scores:
Pred. No.: 8,95e-05 Length: 6414
Score: 151.00 Matches: 93
Percent Similarity: 36.57% Conservative: 88
Best Local Similarity: 18.79% Mismatches: 200
Query Match: 6.37% Indels: 114
DB: Gaps: 14

US-09-889-314-2 (1-496) x US-09-134-001C-1626 (1-6414)
QY 5 SerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeuThr 24
Db 2470 AGTGTAAAGCGAGTCATCATTTACACACAGTAAGAGCTCTCGAGTCCGTTAAGTGG 2529
QY 25 SerThrProGlnGlyValProGlnAsnLysLysSerGlyAsnGluThrLysGlnIle 44
Db 2530 TCGAGCTGCAATAGTACGTCATCATCAGGACGAGCATCGACCAAGCAAGCAAGTAACATTA 2589
QY 45 GlnGlnThrArgGlnGlyLysAsnThrGlnMetGluSerAspAlaThrIleAlaGlyAla 64
Db 2590 TCGAGATCGACC-----AGTACGTCGACATCGATGATGATGCATCGACATCGACCA 2637
QY 65 SerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGlyAl 84
Db 2638 ACAAAAGATGATCATCAACGATAAAGTACGTATTAAGTAATCCACCAAGCAAGCGTT 2697
QY 85 AlaAlaGlyLysGluSerSerGlnLysLysAlaGlyAlaAspThrGlyValSerGly 104
Db 2698 TCGAGTTTCGCAAGCGGTGCACAGTCAGAAAGTGGGTCAACATCAACAAAGTGTGAGTGC 2757
QY 105 AlaAlaAlaThrThrAlaSer---AsnThrAlaThrLysLysIleMetGlnThrSerIle 123
Db 2758 TCAACACAGTACAGCATTTTCAGATCTCACGACGACATCGCGCTGTGATAGCGCATCAATA 2817
QY 124 GlnGluAlaA----- 126
Db 2818 AAAGCGAGGAAATCGACCAAGCAAGTAATATTCAGATTCAGTGCATCGACG 2877
QY 127 -----SerLysSerMetGlnLysThrLeuGlnLysSerLeuSerAlaIleGln 144
Db 2878 TCTGATAGCGCATCAACGTCACAGCGCCTAATGACTCCAAATAGCGCAAGTACGTCATTA 2937
QY 145 MetLysGluValGlnAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlySerAla 164
Db 2938 AGTAAATCGACCAAGCAAGCGCTTCAGATTCACAAAGTACATCGACATCCGACGATGCG 2997
QY 165 LysLeuGluThrProGluLeuProLysProGlyValThrProArgSerGlnValIleGlu 184
Db 2998 TCCACATCAACAGTGTAG-----AGTGTGTCAGCACT 3030
QY 185 IleGlyLeuAlaLeuAlaLysAlaIleGlnThr---LeuGlyGluAlaThrLysSerAla 203
Db 3031 GCAAGTACGTCATTAAAGTAAATCGACAACTACACAGCGTTTCGATTCACACAGAGTACGTCG 3090
QY 204 LeuSerAsnThrLysSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlu 223

```

```

MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
PCT-US93-03993-1

Alignment Scores:
Pred. No.: 7.3e-05 Length: 6530
Score: 152.00 Matches: 79
Percent Similarity: 33.33% Conservative: 28
Best Local Similarity: 24.61% Mismatches: 162
Query Match: 6.42% Indels: 52
DB: 5 Gaps: 7

US-09-889-314-2 (1-496) x PCT-US93-03993-1 (1-6530)
QY 59 AlaThrIleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThr 78
    ||||| ||||| ||||| ||||| |||||
DB 4067 GCCACCGAGAGAGCGGCTCCGACGCTGGAGCCACCGCTCCGCTCCGGAGTACCTCC 4008
QY 79 AlaProGlnGlnGlyAlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAla 98
    ||||| ||||| ||||| ||||| |||||
DB 4007 TCCGACACCGACCGCGCTCCGAGAGCCACTCCGAGCC-----ACCGCGCT 3954
QY 99 AspThrLysLysSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAla 118
    ||||| ||||| ||||| ||||| |||||
DB 3953 GCCCGCGGAAGAGCGCGCCACAGCCACCGCGCTCCGAGTACCTCCACCGCAGCC 3894
QY 119 MetGlnThrSerIleGlnGluAlaSerLysSerMetGlnSerThrLysGlnSerLeuGln 138
    ||||| ||||| ||||| ||||| |||||
DB 3893 ACCGCTCTGCGCCACCGACCGCTGGAGTAGACCCACCGCGCCACCGCGCAGAGA 3834
QY 139 SerLeuSerAlaAlaGlnMetLysGlnAlaValAlaValAlaAlaAlaSerGly 158
    ||||| ||||| ||||| ||||| |||||
DB 3833 GCCACCTCCACAGCTACACCTCTCCGAGAGCCCGCTCCGAGTACCTCCGAGGAGC 3774
QY 159 LysSerSerGlySerLysLeuGluThrProGlnLeuProLysPro-----GlyVal 176
    ||||| ||||| ||||| ||||| |||||
DB 3773 GCAGAGAGTCTGACTGCTCTCTAGAGAGTAATAGCCCTCCGCGCCGCGCAGAGAGCC 3714
QY 177 ThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeu 196
    ||||| ||||| ||||| ||||| |||||
DB 3713 GCCCGCGAGCTGCCACCGCCACCGACCGAATAGCGCGCGCGCTCCGCGCAACT 3654
QY 197 GlyLysAlaThrLysSerAlaLeuSer-----AsnTyrAlaSer 209
    ||||| ||||| ||||| ||||| |||||
DB 3653 GGACCCACCGACAGAGAGCGCTCCGACGTAAGCTCCGCGCCCGGAGTACGCGCC 3594
QY 210 ThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnLysGlnAlaIleLysIle 229
    ||||| ||||| ||||| ||||| |||||
DB 3593 ACCGACGCTGAGGCTCCGCGCGGAGTAGCCCGCGCGGAGCGAGCC----- 3543
QY 230 AspLysGluArgGlnGluTyrGlnGluMetLysAlaAlaGlnGlnLysSerLysAspLeu 249
    ||||| ||||| ||||| ||||| |||||
DB 3543 ----- 3543
QY 250 GluGlyThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrVal 269
    ||||| ||||| ||||| ||||| |||||
DB 3542 -----ACCGCTAAGACCGCTCCGAGCTGCGGCTGATCTGTTGGGA 3498
QY 270 IleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGly 289
    ||||| ||||| ||||| ||||| |||||
DB 3497 CGATCCGCGCGCGCGCGCC-----GGAGAACACACTCCGCGAGCTGCGCC 3450
QY 290 AlaAlaValGlyAlaAlaAlaAlaGlyLysAlaAlaGlyAlaAlaAlaAlaThrVal 309
    ||||| ||||| ||||| ||||| |||||
DB 3449 GCCGCGACCGGAAATGAGCGCGCGCGAGCTGATCACCACCGCTGAGTACCC 3390
QY 310 AlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAla 329
    ||||| ||||| ||||| ||||| |||||
DB 3389 GCCCGCACAGTAGAGCGCGCTGCTCCACCGCC-----GCCGGAATACCGCGC 3339
QY 330 ValArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPhe 349
    ||||| ||||| ||||| ||||| |||||

```

```

DB 3338 TCCGAGCTGGAGCCACCGCGCCACAGAGTACTTCCGCCCTCCGGA----- 3291
QY 350 IleLysThrLeuValLysAlaIleAlaLysAlaIleSerLysGlyIleSerLysValPhe 369
    ||||| ||||| ||||| ||||| |||||
DB 3290 -----ACCGCGCGCGAGCCACCGAGCTGCGCGCTCCAGAGAGCTCCGAGTAGGA 3237
QY 370 Ala 370
    |||||
DB 3236 GCC 3234

RESULT 33
US-09-134-001C-1515/c
; Sequence 1515, Application US/09134001C
; Patent No. 6380370
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCC
; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: GRC-007
; CURRENT APPLICATION NUMBER: US/09/134,001C
; PRIOR FILING DATE: 1998-08-13
; PRIOR APPLICATION NUMBER: US 60/064,964
; PRIOR FILING DATE: 1997-11-08
; PRIOR APPLICATION NUMBER: US 60/055,779
; PRIOR FILING DATE: 1997-08-14
; NUMBER OF SEQ ID NOS: 5674
; SEQ ID NO 1515
; LENGTH: 2835
; TYPE: DNA
; ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1515

Alignment Scores:
Pred. No.: 2.38e-05 Length: 2835
Score: 151.00 Matches: 93
Percent Similarity: 36.57% Conservative: 88
Best Local Similarity: 18.79% Mismatches: 200
Query Match: 6.37% Indels: 114
DB: 4 Gaps: 14

US-09-889-314-2 (1-496) x US-09-134-001C-1515 (1-2835)
QY 5 SerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeuThr 24
    ||||| ||||| ||||| ||||| |||||
DB 2779 AGTGAAGAAGGAGTCAATCATTTATCAACAGTAGTAGTCTCCAGATCGTTAAGTCG 2720
QY 25 SerThrProGlnGlyValProGlnAspLysLeuSerGlyAsnGlnThrLysGlnIle 44
    ||||| ||||| ||||| ||||| |||||
DB 2719 TCGAGTCAAGTCAATGCTCAATCCAGCGCAATCAGCACACACAAATTA 2660
QY 45 GlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGlyAla 64
    ||||| ||||| ||||| ||||| |||||
DB 2659 TCAGATACGCG-----AGTACGTCGACATCTGATGATGATGATGATGATGATGAT 2612
QY 65 SerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGlyVal 84
    ||||| ||||| ||||| ||||| |||||
DB 2611 AGAAAGTAGTCAACAGATAAAGTACGATTAAGTGAATGACACAAAGCGGTT 2552
QY 85 AlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThrGlyValSerGly 104
    ||||| ||||| ||||| ||||| |||||
DB 2551 TCAGATTCAAGCAAGCTGTCACAGAGTCAAGAGTCAACATCAACAAAGTGTGCGC 2492
QY 105 AlaAlaAlaThrThrAlaSer-----AsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
    ||||| ||||| ||||| ||||| |||||
DB 2491 TCACACATCAAGATTTTCAGACTCAACAGACATGAGTGTGATGAGCGCATCATTA 2432
QY 124 GlnGluAla----- 126
    ||||| ||||| ||||| ||||| |||||
DB 2431 AAAGCGACGATCAAGCAAGCAAGTAATATTATGATGATGATGATGATGATGATGATG 2372
QY 127 -----SerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAlaGln 144
    ||||| ||||| ||||| ||||| |||||
DB 2371 TCTGATAGCGCATCAACGTCACAGAGCGTAAAGTGAATGATCTCAATATGCGCAAGTACGTCATTA 2312

```

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/458,240
FILING DATE:
CLASSIFICATION:
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/08/146,930
FILING DATE:
APPLICATION NUMBER: 07/876,286
FILING DATE: April 30, 1992
APPLICATION NUMBER: No. 6143727 yet assigned (204/132)
FILING DATE: October 29, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION NUMBER: 32,327
REFERENCE/DOCKET NUMBER: 204/152
TELECOMMUNICATION INFORMATION:
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6530 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-458-240-1

Alignment Scores:
Pred. No.: 7.3e-05 Length: 6530
Score: 152.00 Matches: 79
Percent Similarity: 33.33% Conservative: 28
Best Local Similarity: 24.61% Mismatches: 162
Query Match: 6.42% Indels: 52
DB: Gaps: 7

US-09-889-314-2 (1-496) x US-08-458-240-1 (1-6530)
QY 59 AlaThrIleAlaGlyAlaSerGlyLysAspLysThrSerThrThrLysThrGluThr 78
DB 4067 GCCACACAGAGAGCCGCTCCGACAGTGGAGCCACCGCTCCGCTCCGAGTAGCTCC 4008
QY 79 AlaProGlnGlnGlyValAlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAla 98
DB 4007 TCCGACGACGACGCGCTCCGCGAAGAGCCACCTCCGAGCC-----ACGCGCGCT 3954
QY 99 AspThrGlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAla 118
DB 3953 GCCCCGGAAGAGCGCCGCGCACAGCCAGCCGCGCTCCGAGTAGCTCCACCGCAGCC 3894
QY 119 MetGlnThrSerIleGlnGluLysSerLysSerMetGlnSerThrLysLysLeuGln 138
DB 3893 ACCGCTGTCGCGCCACACCGCTGGAGTAGAGCCAGCCGCGCCAGCAGAGAGA 3834
QY 139 SerLeuSerAlaAlaGlnMetLysGluValGluAlaValAlaValAlaLeuSerGly 158
DB 3833 GCCACTCCACAGCTACACCTCTCCGGAAGAGCCCTCCGTACTCTGCTGGGGGGC 3774
QY 159 LysSerSerGlySerAlaLysLeuGlnThrProGlnLeuProLysPro-----GlyVal 176
DB 3773 GCAGAGAGTCTACTGCTGCTGAGAGAGTAGAATAGCCCTCCGCGCCAGAGAGAGCC 3714
QY 177 ThrProArgSerGlnValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeu 196
DB 3713 GCCGCGGACGCTGCCACCGCCACACCGGAATAGCCGCGCGCTCCGCGCCGCAACT 3654
QY 197 GlyGluAlaThrLysSerAlaLeuSer-----AsnThrAlaSer 209
DB 3653 GGAGCGACACACAGAGAGCCGCTCCGAGTAGAGCTCCGCGCCGAGAGAGAGCGGC 3594
QY 210 ThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlnLysGlnAlaIleLysIle 229
DB 3593 ACCGACGCTGAGGCTCCGCGCGAGTAGAGCGCGCGCAGCTGAGAGC----- 3543
```

```

QY 230 AspLysGluArgGluGluTyrGlnGluMetLysAlaAlaGluGlnLysSerLysAspLeu 249
DB 3543 ----- 3543
QY 250 GluGlyThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrVal 269
DB 3542 ----- 3542
QY 270 IleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaIleGly 289
DB 3497 CGATCCGCGCGCGCGCCAGC-----GGAGAACCCCTCCGACGCTGTCGCC 3450
QY 290 AlaAlaValAlaGlyAlaAlaAlaAlaGlyAlaAlaGlyAlaAlaAlaAlaThrThrVal 309
DB 3449 GCCGCGACCGGAATAGCCGCGCGCGCGCGCGCGAGTGCATCCAGCCGACCGAGTAGCC 3390
QY 310 AlaThrGlnIleThrValAlaGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAla 329
DB 3389 GCCGCGACAGTAGAGCGCGCGCTGCTTCACCGCC-----GCCGGAATACCGCC 3339
QY 330 ValArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPhe 349
DB 3338 TCCGACGCTGAGCGACCGCGCGCGCACAGAGTAGCTGCCGCGCGCA----- 3291
QY 350 IleLysThrLeuValLysAlaIleLysAlaLysAlaIleSerLysGlyIleSerLysValPhe 369
DB 3290 -----ACGCGCGCGCGCACCGCGAGCTGCCGCTCCAGAGGAGCGCTCCGAGTAGGA 3237
QY 370 Ala 370
DB 3236 GCC 3234

RESULT 32
PCT-US93-03993-1/c
GENERAL INFORMATION:
SEQUENCE 1, Application PC/TUS9303993
APPLICANT: Roop, Dennis R.
APPLICANT: Rothnagel, Joseph A.
APPLICANT: Greenhalgh, David A.
TITLE OF INVENTION: CONSTITUTIVE AND INDUCIBLE EPIDERMAL
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESS: Fulbright & Jaworski
STREET: 1301 McKinney, Suite 5100
CITY: Houston
STATE: Texas
COUNTRY: U.S.A.
ZIP: 77010-3095
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US93/03993
FILING DATE: 19930428
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Paul, Thomas D.
REGISTRATION NUMBER: 32,714
REFERENCE/DOCKET NUMBER: D-5405
TELECOMMUNICATION INFORMATION:
TELEPHONE: 713/651-5325
TELEFAX: 713/651-5246
TELEX: 762829
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6530 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: double
TOPOLOGY: linear
```

```

APPLICANT: Rothnagel, Joseph A.
APPLICANT: Greenhalgh, David A.
TITLE OF INVENTION: SPECIFIC EXPRESSION VECTORS
TITLE OF INVENTION: AND METHODS OF USE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: LYON & LYON
STREET: 611 West Sixth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90017
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: IBM MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/146,930
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA: including application
PRIOR APPLICATION DATA: described below: two
APPLICATION NUMBER: 07/876,286
FILING DATE: April 30, 1992
APPLICATION NUMBER: No. 5958764 yet assigned (204/132)
FILING DATE: October 29, 1993
ATTORNEY/AGENT INFORMATION:
NAME: Warburg, Richard J.
REGISTRATION/DOCKET NUMBER: 32,327
TELEPHONE: (213) 489-1600
TELEFAX: (213) 955-0440
TELEX: 67-3510
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 6530 bases
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-146-930-1

Alignment Scores:
Pred. NO.: 7.3e-05 Length: 6530
Score: 152.00 Matches: 79
Percent Similarity: 33.33% Conservative: 28
Best local Similarity: 24.61% Mismatches: 162
Query Match: 6.42% Indels: 52
Gaps: 7
DB: 2

US-09-889-314-2 (1-496) x US-08-146-930-1 (1-6530)
QY 59 AlathrllealaglyAlasercllyAspylrhserthrlysthrlnuThr 78
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4067 GCCACCAAGAGAGCGCGCTCGCAGCTGAGACCCACCGCTCCGCGAGTAGCCTCC 4008
QY 79 AlAproglngllyAlAAlaAlaglyllySerserSergllySAlaAla 98
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 4007 TCCGACGACCGCGCGCTCGCGAGAGCCACCTCCGAGCC-----ACCGCGCT 3954
QY 99 AspThrglyValserglyAlAAlaAlaAlaAlaAlaAlaAlaAlaAla 118
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3953 GCCCGCGAAGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3894
QY 119 MetglInThSerlIeGlInuAlaSerlySserMetgluSerThrleuGl 138
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3893 ACCGCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3834
QY 139 SerleuSerAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 158
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

DB 3833 GCCACCTCCACACCTACACACCTCCCGAAGACCCCTCCGTAAGCTCTGCTGGGGGCG 3774
QY 159 LyserserSerglyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 176
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3773 GCAGGAGCTGTAGCTGTCTGTGAGAGAGTAGTAATAGCCCTCCGCGCGAGAGAGCC 3714
QY 177 ThrProArSerglyValIleGlInuAlaAlaAlaAlaAlaAlaAlaAla 196
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3713 GCCCGCGCGAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3654
QY 197 GlylAlaAlaThrlySserAlaAlaSer-----AsnlyAlaSer 209
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3653 GAGCCACACACACAGAGAGCGCGCTCCGAGTAGAGCTCCCGCGCGAGTAGAGCGCC 3594
QY 210 ThrGlAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 229
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3593 ACCGAGCTGAGAGCTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3543
QY 230 AspLySgluArggluTyrGlInuAlaAlaAlaAlaAlaAlaAlaAla 249
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3543 ----- 3543
QY 250 GluGlyThrMetAspThrValAsnThrValMetIleAlaAlaAlaAla 269
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3542 -----ACCGTAGAACCGCTCCGAGCTGTGAGAGTAGAGTAGAGTAGAG 3498
QY 270 IleSerIleValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 289
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3497 CGATCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3450
QY 290 AlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 309
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3449 GCCCGACCGAGATAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3390
QY 310 AlathrlInleThrValAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 329
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3389 GCCCGCACCTAGAGCGCGCGCTGCTCCACCGCG-----GCCGAGTAGCGCGC 3339
QY 330 ValArgInAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 349
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3338 TCCGACGCTGAGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3291
QY 350 IleYsThrleuVallyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 369
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 3290 -----ACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3237
QY 370 Ala 370
|||||
DB 3236 GCC 3234

RESULT 31
US-08-458-240-1/c
; Sequence 1, Application US/08458240
; Patent No. 6143727
GENERAL INFORMATION:
APPLICANT: Roop, Dennis R.
APPLICANT: Rothnagel, Joseph A.
APPLICANT: Greenhalgh, David A.
TITLE OF INVENTION: SPECIFIC EXPRESSION VECTORS
TITLE OF INVENTION: AND METHODS OF USE
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESS:
ADDRESSEE: LYON & LYON
STREET: 611 West Sixth Street
CITY: Los Angeles
STATE: California
COUNTRY: U.S.A.
ZIP: 90017
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: IBM MS-DOS (Version 5.0)
SOFTWARE: WordPerfect (Version 5.1)

```



```

US-08-591-079-5
US-RSULT 29
Sequence 5, Application US/08591079
Patent No. 5972899
GENERAL INFORMATION:
APPLICANT: Zychlinskiy, Arturo
APPLICANT: Chen, Yajing
TITLE OF INVENTION: Apoptosis Induced by Shigella Ipab
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,079
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Livanat, Shmuel
REGISTRATION NUMBER: 33,949
REFERENCE/DOCKET NUMBER: 15661-20017.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0764
TELEX: 90-4030 MRSNFOERSMH
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3106 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Yersinia pseudotuberculosis
STRAIN: ypiii(pilB1)
FEATURE:
NAME/KEY: CDS
LOCATION: 12..1217
OTHER INFORMATION: /gene= "yopB"
US-08-591-079-5

Alignment Scores:
Pred. No.: 1 95%-05 length: 3106
Percent Similarity: 152.50 Matches: 93
Best Local Similarity: 37.53% Conservative: 56
Mismatch: 23.43% Mismatches: 163
Query Match: 6.44% Indels: 85
DB: 2 Caps: 17

US-09-889-314-2 (1-496) x US-08-591-079-5 (1-3106)
QY 152 VALVALAALAALeuSerGlyLySerSerGlySerAlaLysLeuGluIurHProGluLeu 1711
||| ||| |||::: ||| :::: ||| :::: |||
Db 111 GTCGGCGAGAACGTGAAGATATAAATATGACGGGTGAGTCT---CACGGCGTACACTC 1678
QY 172 ProGlyProGlyValThrProArgSerGluValIleGlu---IleGlyLeuAlaLeuAla 1900
||| ||| |||::: ||| ||| :::: |||
Db 168 CTGCACCACTAGACAGTGCTGTCGCCACGCAAGTTACTGTAGAAGACACAACAGCAAGAGTCACT 227
QY 191 LysAlaIleGln-----ThrLeuGlyGluAlaThrLysSerAlaLeuSerAsnTyraLa 2008
||| |||::: ||| ||| ||| :::: ||| |||
Db 228 AATATTATTGAGTGCGTACCCGCGGGCGGCGCAGAGATCTCAACTGATATCAATTAATTAATGTT 287
QY 209 Ser-----ThrGlnAlaGlnAlaAsp-----GlnThrAsnLys 2199

```

```
|||||
Db 3942545 CGCTCCCGCCGCGGCGGTGATGCGCTACCGCCAGC---ACCGCGGTTCCACCGT 3942489
|||
Oy 454 eaLaserLySglnThrGlyLuserAsnGlnuMetThrGlnLysAlaThrLysLeuGly-- 473
:|||||: |||: |||: |||
Db 3942488 TGGCGCGCGCGCGCGGTCCTCCGATGCGCCCGCCAGCGCGGTGCGCGGAGA 3942429
Oy 474 -----AlaGlnLeuLysAlaTyrAlaAlaIleSerGlyAlaIleAlaG 489
|||
Db 3942428 TGGCCAGCGCAGCAGCGCGCGCATTCGCGCGCATTCGCCAGCATACCGTTGG 3942369
Oy 489 yAla 490
|||
Db 3942368 CGCC 3942365

RESULT 28
US-09-103-840A-2/c
; Sequence 2, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; TITLE OF INVENTION: TUBERCULOSIS
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: Patentln Ver. 2.1
; SEQ ID NO 2
; LENGTH: 4403765
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; FEATURE:
; OTHER INFORMATION: CDC 1551
; OTHER INFORMATION: "n" bases at various positions throughout the sequence
; OTHER INFORMATION: represent a, t, c or g
US-09-103-840A-2

Alignment Scores:
Pred. No.: 1.4 Length: 4403765
Score: 155.00 Matches: 108
Percent Similarity: 32.48% Conservative: 45
Best Local Similarity: 22.93% Mismatches: 206
Query Match: 6.54% Indels: 112
Db: 4 Gaps: 13

US-09-889-314-2 (1-496) x US-09-103-840A-2 (1-4403765)
Oy 59 AlaThrIleAlaGlyAlaSerGlyLysAspLysThrSerThrThrLysThrGlnThr 78
|||
Db 3938056 GCCGAGGCGCGCGGTGCGCGGT-----GCCGCGCGTCCCGCGCGCGCGCC 3938009
|||
Oy 79 AlaProGlnGlnGlyAlaAlaGlyLysLuserSerGlnLysAlaGlyAla 98
|||
Db 3938008 GTCCCGCCCGCGCGGT---GCCGCGGT-----GCCGCGATGCCCGCATTCGCCGCC 3937958
|||
Oy 99 AspThrGlyValSerGlyAlaAla-----AlaThrThrAlaSerAsnThrAlaThrLys 116
|||
Db 3937957 ACCCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3937898
|||
Oy 117 IleAlaMetGlnThrSerIleGlnLysLuserSerMetGlnLysThrLeuGlnLuser 136
|||
Db 3937897 GCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3937838
|||
Oy 137 LeuGlnSerLeuSerAlaAlaGlnMetLysGlnAlaAlaValAlaVal----- 153
|||
Db 3937837 GCCCGCGGTGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3937778
|||
Oy 154 -----AlaAlaLeuSerGlyLys 159
|||
```

```
Db 3937777 GCGGTGCGCGCGCGATGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCG 3937718
Oy 160 SerSerGlySerAlaLysLeuGlnThrProGlnLeuProLysProGly----- 175
|||
Db 3937717 GCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3937658
|||
Oy 176 -----ValThrProGlnSerGlnValIleGlnIleGlyLeuAlaLeu 189
|||
Db 3937657 GTCCCGCGCGTTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3937598
|||
Oy 190 AlaLysAlaIleGlnThrLeuGlnGlyAlaThrLysSerAlaLeuSerAsnThrAlaSer 209
|||
Db 3937597 TCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3937538
|||
Oy 210 ThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlnLysGlnAlaIleLysIle 229
|||
Db 3937537 ACCGCGTACGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3937481
|||
Oy 230 AspLysGlnThrGlnGlnThrGlnLysGlnLysAlaAlaGlnLysSerLysAspLeu 249
|||
Db 3937481 ----- 3937481
|||
Oy 250 GlnGlyThrMetAspThrValAsnThrValMetIleAlaAlaSerValAlaIleThrVal 269
|||
Db 3937480 -----GCCGCGACCGCG 3937469
|||
Oy 270 IleSerIleValAlaAlaIlePheThrLysGlyAlaGlyLeuAlaGlyLeuAlaGly 289
|||
Db 3937468 GTCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3937409
|||
Oy 290 AlaAlaValGlyAlaAlaAlaAla-----GlyGly 299
|||
Db 3937408 GCCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3937349
|||
Oy 300 AlaAlaGlyAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaValAl 319
|||
Db 3937348 GATGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3937289
|||
Oy 320 GlnAlaValLysGlnAlaValIleThrAlaValAlaThrGlnAlaIleThrAlaIleLys 339
|||
Db 3937288 GCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3937229
|||
Oy 340 AlaAlaValLysSerGlyLysAlaPheIleLysThrLeuValLysAlaIleAlaLys 359
|||
Db 3937228 AGCGCG-----GCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3937175
|||
Oy 360 AlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsn 379
|||
Db 3937174 GGTGCCATTCGCGGATTCAG----- 3937154
|||
Oy 380 PheProLysLeuSerLysValIleSerSerLeuThrSerLysTrpValThrValGlyVal 399
|||
Db 3937153 -----GTTGCCGCGCGTTCACCGCGCATTCGCCACCGTTCGCGCGCTC 3937115
|||
Oy 400 GlyValValValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGln 419
|||
Db 3937114 GCCCGCGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3937064
|||
Oy 420 MetGlnGlnAsnValAlaGlnPheGlnLysGlnValGlyLysLeuGlnAlaAlaAlaAsp 439
|||
Db 3937063 ACCACCGCGCGCGTTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3937004
|||
Oy 440 MetIleSerMetPheThrGlnPheThrPheGlnGlnAlaSerLysIleAlaSerLysGlnThr 459
|||
Db 3937003 GGTGCG-----GTCCGCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCG 3936953
|||
Oy 460 GlyLysSerAsnGlnMetThrGlnLysAlaThrLysLeuGlnAlaGlnIleLeuLysAla 479
|||
Db 3936952 CTGACCGCGCGGTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3936896
|||
Oy 480 TyrAlaAlaIleSerGlyAlaIleAlaGlyAla 490
|||
Db 3936895 GCGCGCGCGCGTTCACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3936863
|||
```

```

OY 309 1A1a-----ThGlnIleThrValGlnAlaValAlaGlnAlaValLysGI 324
    |||
    |||
    |||
Db 875 TGCGCTGAGTCAGCATCAGCAGAGTACGCTCAGCAAGACATCAGCTTCGATGC 934
OY 324 nAlaValIleThrAlaValAlaGlnAlaIleThrAlaValLysAlaValLysSe 344
    |||
    |||
    |||
Db 935 TGCAATCAACAGCGTGCAGCTCAGCTCAGCATCAGCAGCCGCTCAGCTCAGCAAGTAC 991
OY 344 rGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLysGI 364
    |||
    |||
    |||
Db 992 CAGTCCGTCAGCTCAGCAAGTACGATGCTTCAGCTCAGCTCAGCAAGTGCCTGCGG 1051
OY 364 Y 364
    |
Db 1052 C 1052

RESULT 27
US-09-103-840A-1/c
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103.840A
; CURRENT FILING DATE: 1998-06-24
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1

Alignment Scores:
Pred. No.: 1.25 Length: 4411529
Score: 155.50 Matches: 126
Percent Similarity: 31.37% Conservative: 44
Best Local Similarity: 23.25% Mismatches: 193
Query Match: 6.56 Indels: 180
DB: 4 Gaps: 18

US-09-889-314-2 (1-496) x US-09-103-840A-1 (1-4411529)
OY 59 AlAThrlleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThr 78
    |||
    |||
    |||
Db 3943780 GCCCGCGTGGCGGTCGCTC-----GTTGCCGCTT 3943754
OY 79 AlAProGlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAla 98
    |||
    |||
    |||
Db 3943753 GCCCGCGTGGCGGTCGCTC---GCCGCGCTGCCGCGCAGCGCGCTGGCGGTCGCC 3943697
OY 99 AspThrGlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAla 118
    |||
    |||
    |||
Db 3943696 GCCCGCGTGGCGGTCGCTCAGCGCGCGCGCGCGCGCTTACACACACCGCTGGCGGCC 3943637
OY 119 MetGlnThrSerIleGluGlnLysSerMetGluSerThrLeuGlnSerLeuGln 138
    |||
    |||
    |||
Db 3943636 -----ACGCGCGCGCGTGCCTTGGC 3943616
OY 139 SerLeuSerAlaAlaGlnMetLysGluValGlu-----AlaValValVal 153
    |||
    |||
    |||
Db 3943615 GGAATTACGCTTGGCGCGCTTACCGCATTCGCGCGCGCGCGCGCGCTGCC 3943556
OY 154 AlaAlaLeuSerGlyLysSerSerGlySerAlaLysLeuGlnThr-ProGluLeuProLys 173
    |||
    |||
    |||
Db 3943555 GCCCGTGGCGCGTGGCGCTCAGCTTGCCTGCCCGCGCGCGCGCGCGCGCTTACCGCGC 3943496

```

```

OY 173 s-----ProGlyValThrProArgSerGluValIleGI 184
    |||
    |||
    |||
Db 3943495 TTGTCCACCGGAGCGCGGAGGACCGGACCGCCAGCGGTGGCGGTGCCGCGGTCGC 3943436
OY 184 uIleGlyLeuAlaLeuAlaLys---AlaIleGlnThrLeuGlnGlyGlnAlaThrLysSerAl 203
    |||
    |||
    |||
Db 3943435 -GCCGCGACCGCGCGTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3943377
OY 203 AleuSerSerThrAlaSerThrGlnAlaGlnAlaAspGlnAlaThrLysLysLeuGlnI 223
    |||
    |||
    |||
Db 3943376 CTGTGCGCGCGCGCGCGCGCGCGCTTCCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3943351
OY 223 uLysGlnAlaIleLysIleAspLysGluArgGluGluThrGlnGluMetLysAlaAlaGI 243
    |||
    |||
    |||
Db 3943350 -----GGCGCGCGCGGTGCCG 3943314
OY 243 uGlnLysSerLysAspLeu-----GluGlyThrMetAspThrVa 256
    |||
    |||
    |||
Db 3943313 CGGTGCGCGCGGTGATTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCAC 3943254
OY 256 LAsnThrValMetIleAlaValAlaSerValAlaIleThrValIleSerIleValAla----- 274
    |||
    |||
    |||
Db 3943253 CTGTACACACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3943194
OY 275 -----AlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAl 288
    |||
    |||
    |||
Db 3943193 CCGCGCGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3943134
OY 288 aGlyAlaAlaValAlaGlyAlaAlaAlaAla-----GlyAlaAlaAlaGlyAlaAl 304
    |||
    |||
    |||
Db 3943133 CGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3943074
OY 304 AlaAlaAla----- 306
    |||
    |||
    |||
Db 3943073 CGGTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3943014
OY 307 -----ThrThrValAlaThrGlnIleThrValGlnAl 317
    |||
    |||
    |||
Db 3943013 CGCGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3942954
OY 317 aValVal-----GlnAlaValLysGlnAlaValIleThrAlaValArg----- 331
    |||
    |||
    |||
Db 3942953 CGTTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3942894
OY 332 -----GlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPh 349
    |||
    |||
    |||
Db 3942893 CGCGGTCCGCGCGCGTGGCGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3942834
OY 349 eIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLysGly----- 364
    |||
    |||
    |||
Db 3942833 CGCGACCATGCGCGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3942774
OY 365 ----- 365
    |||
    |||
    |||
Db 3942773 CCCCACGTTGCTGCAGCGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3942714
OY 365 eSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPhe----- 380
    |||
    |||
    |||
Db 3942713 CGGTACGTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3942654
OY 381 -----ProLysLeuSerLysValIleSerSerLeuThrSerLysThr 394
    |||
    |||
    |||
Db 3942653 TTTCGCGCTTCCGCGCAATCCACCGCTTCCGCGCGCGCATTTCCGCTACACCGCGCTC 3942594
OY 394 pValThrValGlyValGlyValValAlaAlaProAlaLeuGlyLysGlyIleMetGI 414
    |||
    |||
    |||
Db 3942593 CGCGACGCGCGCGCGCATGCGCGCGCTTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 3942534
OY 414 nMetGlnLeuSerGluMetGlnGlnAsnValAlaGlnPheGlnLysGluValGlyLysLe 434
    |||
    |||
    |||
Db 3942533 -----GGTACCTC 3942546
OY 434 uGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPheThrGlnGlnAlaSerLysI 454

```

```

Db 397 GGTTCAGATCAACGAGTACG---TCAGCTTCAGGCTCAACCAAGCTTCAGCTCAGCA 453
QY 217 ThrAsnLysLeuGlyLeuGlyLysGlnAlaIleLysIleAspLysGluArgGluIuTyr 236
Db 454 TCACACAGTGGCTCAGCTCA----- 474
QY 237 GlnGluMetLysAlaAlaGlnLysSerLysAspLeuGluGlyThrMetAspThrVal 256
Db 475 GCAAGTATTCAGGCTGTAATCGCATCAACGATGCGCTTCAGTACCATCAACGAGT 534
QY 257 AsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIle 276
Db 535 AGCTCAGCTTCAGACAGCACTCAGCTCT-----GAATCGGCC 573
QY 277 PheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAlaIle 296
Db 574 TCACACAGTGGCTCAGCTCAGCATCAACGATGCGCTTCAGTACCATCAACGAGT 633
QY 297 AlaGlyGlyAlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 316
Db 634 TCAGCTCAGCTCAGCAAGTGGCTCAGCTCAGCAACGATGCTGTA-----TCG 684
QY 317 AlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAla 336
Db 685 GCATCAACCAAGTGGCTCAGCTCAGCATCAACGATGCTGCTCAGCTCAGCATCAACCAAGT 744
QY 337 AlaIleLysAlaAlaValLysSer 344
Db 745 GCCTGCGCTTCAGCTCAGCAACGAGT 768

RESULT 26
US-08-961-527-373
; Sequence 373, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8512
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 373:
SEQUENCE CHARACTERISTICS:
LENGTH: 1068 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-373
Alignment Scores:

```

```

Pred. No.: 1,21e-06 Length: 1068
Score: 157.00 Matches: 89
Percent Similarity: 38.58% Conservative: 58
Best Local Similarity: 23.36% Mismatches: 157
Query Match: 6.63% Indels: 78
DB: 4 Gaps: 10

US-09-889-314-2 (1-496) x US-08-961-527-373 (1-1068)
QY 2 ThrAsnMetSerIleSerSerSerGlyProAspAsnGlnLysAsnIle----- 18
Db 90 ACAGTGGCTGCTGATCAGCATCAACGATGCGCTCAGTACCATCAACGATCAAGTCTTGG 149
QY 19 -----MetSerValIleThrThrThrThrProGlnGlyValProGlnLysAspLysLe 36
Db 150 AATCTGATCAACAGTGGCTGATCAGCATCAACGATGCGCTCAGTACCATCAACGATGAC 209
QY 36 uSerClyAsn---GluThrLysGlnIleGlnIleThrArgGln----- 49
Db 210 CAGTGGCTGCTCAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 269
QY 50 -GlyLysAsnThrGluMetGluSerAspAlaThrIleAlaGlyAlaSerGlyLysAspLys 69
Db 270 CAGCAAGTACAGTGGCTGATCAGCATCAACGATGCGCTCAGTACCATCAACGATGAC 323
QY 69 sThrSerSerThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 89
Db 324 TGCATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 382
QY 89 uSerSerGluSerGluLysAlaGlyAlaAspThrGlyAlaSerGlyLysAspLysThr 109
Db 383 AGCATCAACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 442
QY 109 rAlaSerAsnThrAlaThrThrThrThrThrThrThrThrThrThrThrThrThrThrThrThr 129
Db 443 TGGCTGATGCTTGGCATCAACGATGCGCTCAGTACCATCAACGATGATGATGATGATGATGAT 502
QY 129 rMetGluSerThrLeuGluSerLeuGlnSerLeuSerIleAlaGlnMetLysGluValGln 149
Db 503 AGCTCAACCAAGTGGCTCAGCTCAGCATCAACGATGCGCTCAGTACCATCAACGATGATGAT 550
QY 149 uAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlyLysAlaLysLeuGluThrPr 169
Db 551 AGCAAGTATCTCAGCGTCAATCGCATCAACGATGCGCTCAGTACCATCAACGATGATGATGAT 610
QY 169 oGluLeuProLysProGlyValThrProArgSerGluValIleGluIleGlyLeuAlaIle 189
Db 611 T-----GCATC 616
QY 189 uAlaLysAlaIleGlnThrLeuGlyGlnAlaThrLysSerAlaLeuSerAsnThrAlaSer 209
Db 617 AGCTCAGCAAGCAACGATGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGAT 676
QY 209 rThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGluLysGlnAlaIleLysIle 229
Db 677 AACCAAGTGGCTCAGCTCAGCATCAACGATGCGCTCAGTACCATCAACGATGATGATGATGAT 704
QY 229 eaSpLysGluArgGluGluGluGluGluGluGluGluGluGluGluGluGluGluGluGlu 249
Db 705 -----TCAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 731
QY 249 uGluGlyThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrPr 269
Db 732 -----GCATGCGCTTCAGCAACGATGCGCTCAGTACCATCAACGATGATGATGATGATGAT 754
QY 269 IleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGln 289
Db 755 TGCTCAGCTCAGCAAGTATCTCAGCGTCTGATGATGATGATGATGATGATGATGATGATGATGAT 814
QY 289 yAlaAlaValAlaGlyAlaAlaAlaGlyAlaGlyAlaGlyAlaAlaAlaAlaAlaAla 309
Db 815 AGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 874

```

DB: 4 Gaps: 3

US-09-889-314-2 (1-496) x US-08-961-527-370 (1-750)

OY 57 SeraspalarthrIlealaglyAlaserglyLysasplysrThrserserThrThrlysr 76
||| ::||| ||| ||||| ||| ::||| |||
Db 19 TCACATCAACAGAGTGGCTCAGCCTCAGCAAGTACTGATGCTACATCAGCATCAAC 78
OY 77 GluThrAlaprogInglInglyAlaIlealaglyLysgluSerSerSerSerSerSer 96
::: ::||| ::||| ::||| ::||| ::|||
Db 79 AGTCATCAAGCCTCAGCAAGTACTGATGCTGCAATCGCATCAACAGATGATCAGCA 138
OY 97 GlyAlaAspThrGlyValSerGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 116
||| ||| ||| ::||| ::||| ::||| ::||| ::|||
Db 139 TCACATCAACAGAGTGGCTCAGCCTCAGCAAGTACTGATGCTACATCAGCATCAAC 198
OY 117 IleAlaMetGlnThrSerIleGluGluAlaSerLysSerMetGlnSerThrLeuGlnSer 136
||| ::||| ::||| ||||| ||| ::|||
Db 199 AGTCGTCGCTTCAGCATCAACAGAGTGGCTCAGCAAGTACTGATGCTGCA 258
OY 137 LeuGlnSerLeuSerAlaAlaAlaMetLysGluValGluAlaValAlaAlaAla 156
||| ||||| ::||| ||| |||||
Db 259 TCGCATCAACAGAGTGGCTCAGCCTCAGCAAGTACTGATGCTGCA 306
OY 157 SerGlyLysSerSerGlySerAlaLysLeuGluThrProGluLeuProLysProGlyVal 176
||| ::||| ||||| ||| |||
Db 307 TCACAGCAACAGAGTGGCTCAGCCTCAGCAAGTACTGATGCTGCA 345
OY 177 ThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeu 196
346 -----GCCTCACGCTCAGCAAGTACTGATGCTGCA 372
OY 197 GlyGluAlaThrLysSerAlaLeuSerAsnThrLysSerThrGlnAlaGlnAlaAspGln 216
||||| ::||| ::||| ||||| ||| |||
Db 373 TCTGATTCGATCAACAGAGTGGCTCAGCCTCAGCAAGTACTGATGCTGATCAGCATCAGCA 432
OY 217 ThrAsnLysLeuGlyLeuGluLysGlnAlaIleLysIleAspLysGluArgGluLys 236
::: ::||| ::||| ::||| ::||| ::|||
Db 433 TCACAGC----- 438
OY 237 GlnGluMetLysAlaAlaGluGlnLysSerLysAspLeuGluGlyThrMetAspThrVal 256
||| ::||| ||| ||||| ||| |||
Db 439 -----AGTCATCGCTTCAGCAAGTACTGATGCTCAGCAAGTACTGATGCTGCA 489
OY 257 AsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIle 276
::: ::||| ::||| ::||| ::||| ::|||
Db 490 GCGTCAGCCTCAGCAAGTACTGATGCTCAGCCTCAGCAAGTACTGATGCTGATCAGCATCAGCA 549
OY 277 PheThrCysGlyAlaGlyLeuAlaAlaGlyLeuAlaAlaGlyAlaAlaValAlaAlaAla 296
::: ||| ::||| ::||| ::||| ::|||
Db 550 GCAATACAGTGGCTCAGCCTCAGCCTCAGCAAGTACTGATGCTGATCAGCAAGTACTGCA 609
OY 297 AlaGlyLysAlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 316
||| ::||| ::||| ::||| ::||| ::|||
Db 610 GCGTTCGATCAAGTACTGATGCTCAGCCTCAGCAAGTACTGATGCTGATCAGCATCAGCA 669
OY 317 AlaValAlaGlnAlaValLysGlnAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 336
||| ::||| ::||| ::||| ::||| ::|||
Db 670 GCAATACAGTGGCTCAGCCTCAGCCTCAGCAAGTACTGATGCTGATCAGCAAGTACTGCA 729
OY 337 AlaIleLysAlaAla 341
||| ::||| ::||| ::||| ::|||
Db 730 GCGTCATCAAGTACTGATGCTCAGCCTCAGCAAGTACTGATGCTGATCAGCATCAGCA 744

RESULT 25
; Sequence 372, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:

```

1 ADDRESS: Human Genome Sciences, Inc.
2 STREET: 9410 Key West Avenue
3 CITY: Rockville
4 STATE: Maryland
5 COUNTRY: USA
6 ZIP: 20850
7
8 COMPUTER READABLE FORM:
9 MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
10 COMPUTER: HP Vectra 486/33
11 OPERATING SYSTEM: MSDOS version 6.2
12 SOFTWARE: ASCII Text
13
14 CURRENT APPLICATION DATA:
15 APPLICATION NUMBER: US/08/961,527
16
17 FILING DATE:
18 CLASSIFICATION: 424
19 PRIOR APPLICATION DATA:
20 APPLICATION NUMBER:
21 FILING DATE:
22
23 ATTORNEY/AGENT INFORMATION:
24 NAME: Brookes, A. Anders
25 REGISTRATION NUMBER: 36,373
26 REFERENCE/DOCKET NUMBER: PB340P1
27 TELECOMMUNICATION INFORMATION:
28 TELEPHONE: (301) 309-8504
29 TELEFAX: (301) 309-8512
30
31 INFORMATION FOR SEQ ID NO: 372:
32
33 SEQUENCE CHARACTERISTICS:
34 LENGTH: 807 base pairs
35 TYPE: nucleic acid
36 STRANDEDNESS: double
37
38 TOPOLOGY: linear
39
40 US-08-961-527-372
41
42
43
44 Alignment Scores:
45
46 Pred. No.: 5 41e-07 length: 807
47 Score: 158.50 Matches: 66
48 Percent Similarity: 38.54% Conservative: 45
49 Best Local Similarity: 22.92% Mismatches: 140
50 Query Match: 6.69% Indels: 37
51 DB: 4 Caps: 5
52
53 US-09-889-314-2 (1-496) x US-08-961-527-372 (1-807)
54
55 QY 57 SerAspAlaThrIleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThr 76
56 ||| ::::||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
57 16 TCAGCATCAACAGACGCTCGCTTCACGATCAACACGATGCTCGCTTCACGATCAAC 75
58 QY 77 GIuThrAlaProGlnGlnGlyAlaAlaAlaSerLysGlnSerSerGlnSerGlnLysAla 96
59 ||| ::::||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
60 76 ACTGCGTCACATTCAGCAAGTACGACGTCGCTTCACGATCAACACGATGCTTCAGCC 135
61 QY 97 GlyAlaSerThrGlyValSerGlyValAlaAlaAlaThrThrAlaSerAsnThrAlaThrLys 116
62 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
63 Db 136 TCAGCATGACACAGTCCCTCGCTTCACGACGACATCAAGCATTCGATTCAGGCTCAAC 195
64 QY 117 IleAlaMetGlnThrSerIleGlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSer 136
65 ||| ::::||| ||| ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
66 Db 196 ACTGCTTCGGCTTCAGCAAGTACGACGTCCTTCACGCTTCACGATCAACACGCTCGCC 255
67 QY 137 LeuGlnSerLeuSerAlaAlaGlnMetLysGlnValGlnAlaValAlaAlaAlaLeu 156
68 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
69 Db 256 TCACGACACACCTCGAGTTCGAAATCGGCTCAACACGACGCTCGGCTCGACGACGACC 315
70 QY 157 SerGlyLysSerSerGlySerAlaLysLeuGlnThrProGlnLeuProLysProGlyVal 176
71 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
72 Db 316 TCAGCTTCGAATCGGCTCA----- 336
73 QY 177 ThrProArgSerGlnValIleGlnIleGlyLeuAlaAlaLeuAlaLysAlaIleGlnThrLeu 196
74 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
75 Db 337 ACACAGCGGCTCGGCTCGACATCAACGAGTCTCGGCTTCAGCAACGACACAGCGCTCG 366
76 QY 197 GlyGlnAlaThrThrLysSerAlaLeuSerAsnThrLysSerThrGlnAlaGlnAlaAspGln 216
77 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 368:
SEQUENCE CHARACTERISTICS:
LENGTH: 941 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-368

Alignment Scores:
Pred. No.: 1,08e-07 Length: 941
Score: 166.50 Matches: 70
Percent Similarity: 39.94% Conservative: 57
Best Local Similarity: 22.01% Mismatches: 180
Query Match: 7.03% Indels: 12
DB: Gaps: 2

US-09-889-314-2 (1-496) x US-08-961-527-368 (1-941)
QY 27 ProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGluThrLysGlnIleGlnGln 46
DB 20 CCTCACACAGTGGCTTCAGCCTCAGCCTCAACCAAGTGGCTTCAGCAGAGTACAGT 79
QY 47 ThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThrIleAlaGlyAlaSerGly 66
DB 80 CGTCACTTCAGCAAGCAAGTGGCTTCAGTCAAGTCAACCAAGTGGCTTCAGCA 139
QY 67 LysAspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGlyAlaAla 86
DB 140 TCACAGAGTGGCTTCAGTCAAGTGGCTTCAGTCAAGTGGCTTCAGTCAAGT 198
QY 87 GlyLysLeuSerGluSerGlnLysAlaGlyAlaAspThrGlyAlaSerGlyAlaAla 106
DB 199 TCAGCATCAGCATCAACCAAGTGGCTTCAGTCAAGTGGCTTCAGTCAAGT 258
QY 107 AlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGluGluAla 126
DB 259 TCACAGAGTGGCTTCAGTCAAGTGGCTTCAGTCAAGTGGCTTCAGTCAAGT 318
QY 127 SerLysSerMetGluSerThrLysLeuSerLeuGlnSerLeuSerAlaAlaGlnMetLys 146
DB 319 TCAGTCTCAGCAAGCAAGTGGCTTCAGTCAAGTGGCTTCAGTCAAGTGGCT 378
QY 147 GluValGluAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlySerAlaLysLeu 166
DB 379 ACTATCTCAGGCTTCGATTCGCAATCAGCAAGTGGCTTCAGTCAAGTGGCT 438
QY 167 GluThrProGluLeuProLysProGlyValThrProArgSerGluValIleGluIleGly 186
DB 439 TCGGCTTCAGCAAGCAAGTGGCTTCAGTCAAGTGGCTTCAGTCAAGT 483
QY 187 LeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGluAlaThrLysSerAlaLeuSerAsn 206
DB 484 ---GCTTCAGCTTCAGCAAGTATCTCAGGCTTCGATTCGATCAACCAAGTGGCT 540
QY 207 TyrAlaSerThrGlnAlaGlnAlaLysGlnThrAsnLysLeuGlnGluLysGlnAla 226
DB 541 TCAGCAAGTACTAGTGCATCAGCATCAGCATCAGCATCAGCATCAGCATCAG 585

```

```

QY 227 IleLysIleAspLysGluArgGlnGluThrGlnGluMetLysAlaAlaGlnGlnLysSer 246
DB 586 GCTTCAGCAAGTACCAAGCGCTTCAGTCAAGCAAGCAAGTGGCTTCAGCAAGT 645
QY 247 LysAspLeuGlnGluThrMetAspThrValAsnThrValMetIleAlaValSerValAla 266
DB 646 ACCAGCGCTTCAGGCTTCAGCAAGTGGCTTCAGTCAAGTGGCTTCAGCAAGT 705
QY 267 IleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeu 286
DB 706 GCTTCAGCTTCAGCAAGTGGCTTCAGTCAAGTGGCTTCAGCAAGTGGCTTCAGCAAGT 765
QY 287 AlaAlaGlyAlaAlaValGlyAlaAlaAlaAlaGlyGlyAlaAlaGlyAlaAlaAlaAla 306
DB 766 ACGATGCTCAGCTTCAGTCAAGTGGCTTCAGTCAAGTGGCTTCAGTCAAGTGGCT 825
QY 307 ThrThrValAlaThrGlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaVal 326
DB 826 GCTTCAGTCAAGCAAGTGGCTTCAGTCAAGTGGCTTCAGTCAAGTGGCTTCAGCAAGT 885
QY 327 IleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSer 344
DB 886 ACAAGTGGCTTCAGTCAAGCAAGTGGCTTCAGTCAAGTGGCTTCAGCAAGTACTAGT 939

```

RESULT 24

```

US-08-961-527-370
Sequence 370, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44M storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 370:
SEQUENCE CHARACTERISTICS:
LENGTH: 750 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
US-08-961-527-370

Alignment Scores:
Pred. No.: 9.41e-08 Length: 750
Score: 165.50 Matches: 62
Percent Similarity: 38.95% Conservative: 49
Best Local Similarity: 21.75% Mismatches: 131
Query Match: 6.99% Indels: 43

```

```

Oy      389 SerleuthrSerlystrPvalThvAlGlyValGlyValValAlaAlaProAlaLeu 408
      ::::||||:::      :::: |||      |||||
Db      12286 GCGCTTACGAACGGGTATTGCTTCTGCGGG-----ACACTGGGTTG 12245
Oy      409 GlylsglylIemEtGlnmetGln---LeuSerGlnmetGln-----GlnAsn 423
      :|||:::||||      |||:::||||:      :|||      |||||
Db      12244 AATANTAGAGCTGGCGCAACTCCAAATGTGTCTTCACGTACTGCTTTCCGGGTGCAAAAC 12185
Oy      424 ValAlaGlnpHeGlnlyGluValGly-----LysLeuGlnAlaAlaAla 438
      ::::      :::: |||      :|||      |||||
Db      12184 TTAACCTGCACAGACTGAAGCTTAGTGAGAGTCGCAAGCTCGAGCTGAATAAGCGGCA 12125
Oy      439 AspmelliesrMetpHeThrGlnpHeTrpGlnAlaSerlylIeAlaSer 456
      ::::      :::: |||      :|||      |||||
Db      12124 AGCGAGTTACAAATACAGCGCAGCTATTTCACAAAGTGTTCACACTGAGATGCC 12071

RESULT 22
US-08-961-527-263
: Sequence 263, Application US/08961527
: Patent No. 6420135
: GENERAL INFORMATION:
: APPLICANT: Charles Kunsch
: TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
: NUMBER OF SEQUENCES: 391
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Human Genome Sciences, Inc.
: STREET: 9410 Key West Avenue
: CITY: Rockville
: STATE: Maryland
: COUNTRY: USA
: ZIP: 20850
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
: COMPUTER: HP Vectra 486/33
: OPERATING SYSTEM: MSDOS version 6.2
: SOFTWARE: ASCII Text
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/961,527
: FILING DATE:
: CLASSIFICATION: 424
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER:
: FILING DATE:
: ATTORNEY/AGENT INFORMATION:
: NAME: Brookes, A. Anders
: REGISTRATION NUMBER: 36,373
: REFERENCE/DOCKET NUMBER: PB340P1
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (301) 309-8504
: TELEFAX: (301) 309-8512
: INFORMATION FOR SEQ ID NO: 263:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 3744 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: US-08-961-527-263

Alignment Scores:
Pred. No.:      9e-07      Length:      3744
Score:          167.00      Matches:      64
Percent Similarity: 42.36%      Conservative: 58
Best Local Similarity: 22.22%      Mismatches: 145
Query Match:      7.05%      Indels:      21
DB:              4      Gaps:      4

US-09-889-314-2 (1-496) x US-08-961-527-263 (1-3744)
Oy      24 ThrSerThrPrpGlnGlyValAlpProGlnGlnAspLysLeuSerGlnGlnGluThrLysGln 43
      ::::||||      |||      :|||      |||||
Db      2924 AGCAGACCACTGACGTACTTTCACAGTCATCACTCTTCACGTAAAGTGGTACTGCAA--- 2980
Oy      44 lIeGlnGlnThrPrpGlnGlnGlyLysAsnThrGlnUmetGlnSerSpAlaThrIleAlaGly 63

```

```

Db      2981 -----:::|||||:::|||||
        AlaSerGlyLysAspLysThrSerSerThrIleThrLeuThrAlaProGlnGlnGly 83
        |||||          |||:::|         |
Db      3014 GCTTACAGCATCAACACAGTGCGCTGGCTTCACGCATCACCAAGTGGTACGTTCACGAAGT 3073
        |||||          |||:::|         |
        ValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyValSer 103
        ::|||          :|||:::|         |
        3074 ACCACTGCTCAGTCCATCCATCAACAAGTCTTCACCTCCACATCGACATCGAATGCGCTCG 3133
        |||||          |||:::|         |
        GlyAlaAlaAlaIleThrAlaSerAsnThrAlaIleThrIleAlaMetGlnTrpSerIle 123
        ::|||          |||:::|         |
        3134 GCTTAGGACAGCATCATGAGATCGAATCAGACCGTCAACAGCGTTCCGCTTCAGCAAGT 3193
        |||||          |||:::|         |
        GluGluAlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeuSerAlaAla 143
        |||||          |||:::|         |
        3194 ACCACTGCTTCACGCTTCACGATCAACACAGCGGCCCTCGGCCTCAGCAAGCACCCTCATTCT 3253
        |||||          |||:::|         |
        GluMetLysGluValGluAlaValValAlaAlaLeuSerGlyLysSerSergLysSer 163
        ::|           |||   |||:::|     |||   |||
        3254 GAATGCGGCTCAACACAGCGCTCGGCCTCAGCACAGCACCCTCAGCTTCGTAATCGCGCTCA 3313
        |||||          |||   |||:::|     |||   |||
        AlaLysLeuGluThrProGluLeuProLysProGlyValIleThrProArgSergLysValIle 183
        164 AlalysLeuGluThrProGluLeuProLysProGlyValIleThrProArgSergLysValIle 183
        3314 ACCAGCGGCTCAGCC-----TCAGCATCAACAGCGTCTCGGCTTCAGCA 3358
        |||||          |||   |||:::|     |||   |||
        GluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGluAlaIleThrLysSerAla 203
        ::|           |||   |||:::|     |||   |||
        3359 AGCACACAGCGCTCGGGTTCACGATCAACAGATGTCGTCAGCTTCAGGTCACACCAAGTCT 3418
        |||||          |||   |||:::|     |||   |||
        LeuSerAsnTyrrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlu 223
        |||   |||   |||   |||   |||   |||   |||
        3419 ---TCAGGCTCAGATCAACAAGTGCACCTCAGCAAGATCTCAGCTCGAATCGG 3475
        |||||          |||   |||:::|     |||   |||
        LysGlnAlaIleLysIleAspLysGluArgGlnLysIleGlnGluMetLysAlaAlaGlu 243
        |||   |||   |||   |||   |||   |||   |||
        3476 CATCAACAGAGTGGCTG----AGTAGATCAACACAGATCAAGTCAAGCTCAG 3523
        |||||          |||   |||:::|     |||   |||
        GluLysSerLysAspLeuGlu---GlyThrMetAspThrValAsnThrValMetIleAlaVala 263
        |||   |||   |||   |||   |||   |||   |||
        3524 CAAAGCACTCAAGCTTCTGAATCGCGCTCACACAGTGGTCACTCACTCAAGCAACAGCC 3583
        |||||          |||   |||:::|     |||   |||
        LysValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLe 283
        |||   |||   |||   |||   |||   |||   |||
        3584 CTCACCTCAGCAAGTACCAAGTCCAGCTCAGCTCGCAAGTGCATGCGGCTCAAC 3643
        ::|           |||   |||:::|     |||   |||
        ValAlaGlyLeuAlaAlaGlyAlaAlaValGlyAlaAlaAlaAlaGlyLysAlaAlaGlyAl 303
        ::|           |||   |||:::|     |||   |||
        3644 CAGTGCATCTGAATGGGATCAACCAAGTGCCTCAGCTCAGCAAGTACTGATGCATCGC 3703
        |||||          |||   |||:::|     |||   |||
        303 aaAlaAlaIleThrThrValAla 310
        Db      3704 TTCAGATCAACCAAGTGCCTCG 3725
        |||||          |||   |||:::|     |||   |||

RESULT 23
US-08-961-527-368
; Sequence 368, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 Inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
```

GENERAL INFORMATION:
APPLICANT: Blattner, Frederick R.
Burland, Valerie
Perna, Nicole T.
Plunkett, Guy
Welch, Rod
TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Charles & Brady
STREET: 1 South Pinckney Street
CITY: Madison
STATE: WI
COUNTRY: US
ZIP: 53701-2113
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 8.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453,702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110,955
FILING DATE: 04-Dec-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296,95017
TELEPHONE: (608) 251-9166
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 261:
SEQUENCE CHARACTERISTICS:
LENGTH: 45325 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 261:
US-09-453-702B-261
Alignment Scores:
Pred. No.: 2,55e-05 Length: 45325
Score: 170.00 Matches: 114
Percent Similarity: 37.55% Conservative: 88
Best Local Similarity: 21.19% Mismatches: 208
Query Match: 7.18% Indels: 128
Gaps: 23
DB: 4
US-09-889-314-2 (1-496) x US-09-453-702B-261 (1-45325)
QY 1 AspThrAsnMetSerIleSerSerSerSerSerProAspAsnGlnLysAsnIle----- 18
Db 13546 GATATGAATGAGGATCTTAAGCGTCAACGACGACGACAAAGATGCTAATCTGTGAT 13487
QY 19 -----MetSerGlnValLeuThrSerThr-----ProGln 28
Db 13486 GCCAAATTTGCTGATGTTAGAGTACACTGTATAGTAAGTGCAGAAACCCAACTTCCTCA 13427
QY 29 GlnVal-----ProGlnInAsp-----Lys 35
Db 13426 GAGCTGATGACTATATATAACATCCACTATGACATTAAGTGAAGTGAATTCGTGAT 13367
QY 36 LeuSerGly----- 38
Db 13366 CTTAGTGTGATATTAAAGCGGTGATCTGCAAAAGTGAAGCGGCTATTTCAGCTAAA 13307
QY 39 ---AsnGlnThrLysGlnIleGlnGlnInThrArgGlnGlnLysAsnThrGlnLysMetGlnSer 57
Db 13306 GCGAATTAACCTGACAAACGATGATTAATATAGCCAGCTCGAAATTCGAAATGTGCAAT 13247

QY 58 AspAlaThrIleAlaIleAlaSerGlyLysAspLysThrSerSerThrThrLysThrGln 77
Db 13246 ACATTAATCTCTTAACGATGACGCTTCTGATGCAATCTCAATATATAGAACTATT 13187
QY 78 ThrAlaProGlnGlnGlnValAlaAlaGlyLys----- 88
Db 13186 TCAGCA-----ATATCCCTTGTTAAATAACGGAATACATGCTTAACGTA 13139
QY 89 ---GluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyValSerGlyAlaAlaAla 107
Db 13138 AATAACGATACCTGCTCTGTATACGCTGAGGCTGTAATACC---GCCTCGGCTACTTCTGT 13082
QY 108 ThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlnLysAlaSer 127
Db 13081 ATTACTCAATCTGAACGGGTTTATCCCTGCATTTACAA-----CTGCTT 13037
QY 128 LysSerMetGluSerThrIleu----- 134
Db 13036 AATTCATGAACCTGCTGACGACGCTGACAGAAAGTACCCCTTTACCGACGCCCGCGCA 12977
QY 135 ---GluSerLeuGlnSerLeuSerAlaAlaGlnMetLysGlnValGlnAlaValVal 153
Db 12976 GGTCACTCATTAAGTACGCGCTCTGCTGCTGAG-----GATGCTCTT 12935
QY 154 AlaAlaLeuSerGlyLysSerSerGlySerAlaLysLeuGlnThrProGlnLysProLys 173
Db 12934 AGTAATTTCTTGTGATGATTAAGTGTGAGGTGACAGTGCACATCGCATGAGAG--GCAGAG 12878
QY 174 ProGlnValThrProArgSerGlnValIleGlnIleGlyLeuAlaAlaValAlaIle 193
Db 12877 CCACACGCGACAGCAAGCTAATCTATCTCTCAGGAAATACCGTTGACCTCAG----- 12824
QY 194 GlnThrLeuGlyGlnAlaThrLysSerAlaLeuSerAsnThrAlaSerThrGlnAlaGln 213
Db 12823 CAATGATGATGATGTGCTACCTGTTATCCCTGCTGATCTCCGCGACAGAAAGTCTCAGT 12764
QY 214 AlaAspGlnThrAsnLysLeuGlyLeuGlnLysGlnAlaIleLysIleAspLysGlnArg 233
Db 12763 CTGAAACACCTCAACGAGATTTATGATGGCGCAAACTAAGCCCTGAGAAATAAACG 12704
QY 234 GlnGlnThrGlnGlnMetLysAlaAlaGlnGlnLysSerLysAspLeuGlnGlyThrMet 253
Db 12703 CAGGAGTATATAAAACAGCTCGAAGAAACAGAAAGCCGAAAGAAATCACAAAAAGT 12644
QY 254 AspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleVal 273
Db 12643 AAAATTGTTGGCCAGGCTTGTGTTGGCTGGCGCTGCAATTAACAGCC-----GTT 12593
QY 274 AlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaValAlaGly 293
Db 12592 GCCGCTGTTTTAAC-----CCAGCACTCTGGCGCTGTTGTTGCTATTGGT 12548
QY 294 AlaAlaAlaAlaGlyAlaAlaIleGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 313
Db 12547 -----GCACACGCAATGGCAGCTGCAACG 12524
QY 314 ThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValAlaGln---- 332
Db 12523 GCAGTGTGATGATATGGGGAAATGCCCTCAGGATTAAGACTGCAGCAGGCTCTT 12464
QY 333 -----AlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPhe 349
Db 12463 GCGGAAATATCTATGCGCCAGCAAGCTTCTGACAGCGCGGCTGGCGGGGTCTTCACTG 12404
QY 350 IleLysThrLeuValLysAlaIleAlaLysAlaIleSerLysGlyIleSerLysValPhe 369
Db 12403 TTATCTAAATTTGCTAAT---GTTCGTAAACAAATTTGCTCAAGCGTGTGTAAGTCTT 12347
QY 370 AlaLysGlyThrGlnMetIleAlaLysAsn---PheProLysLeuSerLysValIleSer 388
Db 12346 GAGAAAGCGGACAGAAAGCTGTTAAAGCGTTTTCGCAAAATTTCCACAGTGGCTGAG 12287


```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/453, 702B
FILING DATE: 03-Dec-1999
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/110, 955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296, 95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ. ID NO: 206:
SEQUENCE CHARACTERISTICS:
LENGTH: 43360
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 206:
US-09-453-702B-206

Alignment Scores:
Pred. No.: 2,38e-05 Length: 43360
Score: 170.00 Matches: 114
Percent Similarity: 37.55% Conservative: 88
Best Local Similarity: 21.19% Mismatches: 208
Query Match: 7.18% Indels: 128
Gaps: 23
DB: 4

US-09-889-314-2 (1-496) x US-09-453-702B-206 (1-43360)
QY 1 Asphthrasmetserliserseisersegllyproapaspnnglnlyasnlle----- 18
Db 12644 GATATGATGAGCGATCTTAAGCGCGACGACCAAAAGATGCTAATCTGTGAT 12585
QY 19 -----Metserglvalleuthrserthr-----Proglin 28
Db 12584 GCCAAATGCGATGATGTCAGAGTAGCACTGATTAAGAAATGCCAAAGCCAACTTCTCA 12525
QY 29 Glyval-----Proglinlasp-----Lys 35
Db 12524 GACGCGATGACTATATTAAGCATCCAGTATGACATTAAGTAACTGTATGCTGAT 12465
QY 36 leusergly----- 38
Db 12464 CTTACTGGTATTTAAGCGCTGCTGATCGCAACAGTCGAAGCGGCTATTTCAGCTAA 12405
QY 39 ---Aaengluthrlysglnlleglnlnthrarglnglylysasnthrghumetgluser 57
Db 12404 GCGAATTAACCTGCAACAGCTAGTAATTAATAGCCAGCTCGAATTCAGCAAAATGCGAAT 12345
QY 58 Aspalatrhlleagllyalaserglylysasplyrthrsersethrthrlysthrglu 77
Db 12344 ACATTAATCTCTTAAGCGAGCTGCTGATGTCATCTCTCAATATATGAACATAT 12285
QY 78 ThrAlaproglnlnglylvalAlaAlaAlaLys----- 88
Db 12284 TCAGCA-----ATATCCCTTGTTAATAATACCGGAGATTAACATGCTTAACGTA 12237
QY 89 ---GluserseglserglinlyslaglylalaasplhrcllylaserlylAlaAla 107
Db 12236 AATTAACGATACCTGCTGTAAAGCTGTGAGTTAAATACC---GCCCTGGGTACTTCTGCT 12180
QY 108 ThrthrAlaSerasnThrAlaThrlyslleAlaMetGlnThrserlleglnlualaser 127
Db 12179 ATTACTCAATCTGAACCGGTTTATCGCTGATTTACAA-----CTGGTT 12135
QY 128 Lyssermetgluserthrleu----- 134
Db 12134 AATTCATGACTGTCGACAGCGCTGACAGAAAGTAGCCCTTTACGAGCCGCCGCGCA 12075
```

```

QY 135 ---GluserleuglnserleuserAlaAlaAlaGlnmetlysglvalgluAlaVal 153
Db 12074 GGTACATCATTAAGACGCGCCCTGCTGCTGAC-----GATGCTCTT 12033
QY 154 AlaAlaLeuserglylyserserglyserAlaLysleuglnthrProglinleuplys 173
Db 12032 AGTAATATTGTTGGTGTATTAAGTAGGAGTTTCAAGTCGCACTGAGAG---GCAGAG 11976
QY 174 ProgluThrProargsergluvalleuglnleuglyleuAlaLeuAlaAlaAla 193
Db 11975 CCACAGCGCACACCTATCCCTATCTCTCAGGTGAATACCTTGACCTCAG----- 11922
QY 194 GlnthrleuglygluAlaThrlysserAlaLeuserAlaThrGlnAlaGln 213
Db 11921 CAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 11862
QY 214 AlaasplnhrasnlyslleuglnlysglnAlaAlaLyslleasplygluarg 233
Db 11861 CTGAATAACTCTCAACGAGATTATATGATGATGATGATGATGATGATGATGATGATGAT 11802
QY 234 GlugluthrlysglnlmetlyslAlaAlaGlnllysserlysaaspleugllythrmet 253
Db 11801 CAGGATTAATAAACAACAGCTGGAAGACACAGAAAGCCGAGAAATCAAAAAAGT 11742
QY 254 AsplhrAlaasnThrValmetlleAlaValaserAlaAlaThrValleuserlleval 273
Db 11741 AAAATGTTGGCCAGGCTCTTGGTGGTGGGCGCGCATTAACAGCC-----CTT 11691
QY 274 AlaAlaThrleuphrlysglylAlaGlyleuAlaGlyleuAlaAlaAlaValaGly 293
Db 11690 GCCGCTGTTTAAAC-----CCAGCAGCTGCGGCTGTTGTTGCTTCAATGTT 11646
QY 294 AlaAlaAlaAlaGlylAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 313
Db 11645 -----GCAACAGCAATGCACTGCAACG 11622
QY 314 ThrValGlnAlaValaGlnAlaValaLysGlnAlaValaThrAlaValaArgln--- 332
Db 11621 GCAGTCGATGAATGAGGGAATAATGCCCTCAGGATTAAGACTGACAGCAGAGCTTTT 11562
QY 333 -----AlaThrAlaAlaAlaLysAlaAlaValLysserglylleLysAlaPhe 349
Db 11561 GCGGGAATATCTATGCGCGCAAGCATCTGACAGCGCGCTGGCGGTGTCCTCACTG 11502
QY 350 lleysthrleuvallyslAlaAlaAlaLysAlaLysleuserlylleuserlyslPhe 369
Db 11501 TTATCTAATTTGTAAT---GTTGCTAACAATAATTTGCTCAAGCGTGTAAAGTCTGTT 11445
QY 370 AlaLysGlyThrGlnmetlleAlaLysasn---PheProLysleuserlyslleuser 388
Db 11444 GAGAAAGCGCGCAAGAACCGCTGTTAAAAACGTTTTTGCAGAAATTTGCAGAGCTGAG 11385
QY 389 SerleuphrserlystrlyslThrValGlylAlaGlylAlaAlaAlaAlaAlaAla 408
Db 11384 GCGCTTAGCAAGCGTATTCCTGCTCCCGG-----ACAACCTCGTTG 11343
QY 409 GlyLysGlyllemerglmetgln---leuserglumetgln-----Glnasn 423
Db 11342 AATTAATGAGCTGCGCAACTCCAAATGTTGTCTCAGTTAGCTCTTTGCGCGTCAAAAC 11283
QY 424 ValAlaGlnPheGlnllysluvalgly-----LysleuglnAlaAlaAla 438
Db 11282 TTAACTCGACAGAGAAAGCTTAGGTGAGACTGCGAAGCTGAGCTGATTAACGCGCA 11223
QY 438 AspmetlleSermetPheThrGlnPheThrpglnlnglnAlaSerlylleAlaSer 456
Db 11222 ACGGATTAACAATAACGAGCGAGCTAATTAACAAGTGTTCACACTGATGTGC 11169
RESULT 21
US-09-453-702B-261/C
; Sequence 261, Application US/09453702B
; Patent No. 6365723
```

```

Db 1372 GACATTTAATCACTAATGCTGTGGCAGATTAATAATTAATTTCTTGTCGACAGGGGTAT 1431
OY 466 -----ThgLnLysAlaThrLysLeuGlyAla 474
Db 1432 GAAGTAATATCCAAACAATTAATTTCCACCCATTTAACCAAGACATTTTATTAGGAGAA 1491
OY 475 GlnLeuLysAlaTyrrAlaAlaIleSerGlyAlaIleAlaGlyAla 490
Db 1492 AGGTATACCTCTGCCACACAGCG--GGAGGAGTGTCCGTTCTGCT 1536

RESULT 19
US-08-961-527-369
; Sequence 369, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8512
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 369:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 869 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-08-961-527-369

Alignment Scores:
Pred. No.: 2.34e-08 Length: 869
Score: 172.50 Matches: 74
Percent Similarity: 41.14% Conservative: 49
Best Local Similarity: 24.75% Mismatches: 138
Query Match: 7.28% Indels: 39
DB: 4 Gaps: 6

US-09-889-314-2 (1-496) x US-08-961-527-369 (1-869)
OY 59 AlaThrIleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnurh 78
Db 6 AGTACTAGTCATCGACTTCAGCATCAACGATGCAATCGCTTCCTCGTCAACCAAGTGA 65
OY 79 AlaProGlnGlnGlyValAlaAlaGlyLysGluSerSerGluSerGlnSerGlnLysAlaGlyAla 98
Db 66 TCAGAGTCAGCAGCAATACAGTCAGTCGCTCAGCTTCGCAATCAACAAGTGCCTCGGCTTCAGCA 125
OY 99 AspThrGlyValSerGlyAlaAlaAlaAlaThrThrAlaSerAsnThrAla---ThrLysIle 117
Db 126 AGCACACAGTCGCTGCGCTTCAGCAAGTACTAGCGCTTCACGCTTCAGCTTCACACCAAGTGC 185

```

```

QY      118 AlameGlnThrSerIle-----GluGluAlaSerLysSerMetGluSerThrLeu 134
      :::: ||||| ||| ||||| ||| ::::
Db      186 TCACCTCAGCAGAGTATCTCAGGCTGTGATGGCTGACCAAGAGCTGGCTCGCTTACGA 245
QY      135 GluSerLeuGlnSerLeuSerAlaAlaGlnMetLysGluValGluAlaValAlaValAla 154
      :::: ||| ||||| ::::
Db      246 AGTACTAGCGCCTCAGCTCAGGCTCAACAAAGTGCATCGGGCTTCAGGCTCAACAGAGCGC 305
QY      155 AlaLeuSerGlyLysSerSerGlySerAlaLysLeuGluThrProGluLeuProLysPro 174
      :::: ||| ||||| ||||| |||
Db      306 TCTGATCGGCATCAACGAGTGGCTCGCTTCAGCAAGTACTAGC----- 350
QY      175 GlyValThrProAlaGlySerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGln 194
      :::: ||| ||| ||| ||| |||
Db      351 -----GCCACAGCCGACGCCGCAACA 371
QY      195 ThrLeuGlyGluAlaThrLysSerAlaLeuSerAsnTyAlaSerThrGlnAlaGlnAla 214
      :::: ||| :::: ||| ||||| ||| |||
Db      372 AGTCATCGGGCTTCAGCATCAACAGAGCGCTCGCTTCACAAAGTACTAGCGCCTCAGCC 431
QY      215 AspGlnThrAsnLysLeuGly-LeuGluLysGlnAlaIleLysIleAspLysGlu----- 232
      :::: ||| ||||| ::::
Db      432 TCAGCCTCACAAGTGCATCGGCTTCACGCTCAGCAAGTGCATGAGTGCATCAACAG 491
QY      233 ----ArgGluGluTyArgGlnLumMetLysAlaAlaGluGlnLysSerLysAspLeuGlu 251
      :|||: ||| ||| :|||
Db      492 AGTGGCTGACCCACAGCAGC-ACATCAGCTTCGATCTGCATCAACACAGTGGCTGAC 550
QY      251 TyrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSe 271
      :||| ||| :||| ||| :|||
Db      551 CTCAGCATCGACAGAGCGCTCAGCTTCAGCAAGTACAGTGCATCAACAGTGCATCA 596
QY      271 rIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaGlyAlaAla 291
      :|||: ||| ||| :||| ||| :|||
Db      597 -----GCCTCAGCGCTGACAGAGTGGCTGCTTCAGCAAGTACAGTGCATCAAGCTC 649
QY      291 aValGlyAlaAlaAlaGlyLysGlyAlaGlyAlaGlyAlaAlaAlaAlaAlaThrThrValAla 310
      :|||: ||| ||| :||| ||| :|||
Db      650 AGCAAGTACCGAGTGGCTGACCTCAGCTCAGCTGCAAGAGTGGCTGACCTCAACAGTGCATC 709
QY      311 -----ThrGlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaVala 326
      :||| ||| :||| ||| |||
Db      710 TGAATCGGCATCAACAGCGTGCAGCTTCAGCTGCAAGTACTAGTGCATCAGCTTCAGCTC 766
QY      326 IleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSer 344
      :|||: ||| ||| :||| ||| :|||
Db      770 AACGAGTGCATGGCTTCAGCATCAACCGAGTGCATCAGAGTACGCAAGTACCACT 824

RESULT 20
US-09-453-702B-206/c
: Sequence 206, Application US/09453702B
: Patent No. 6365723
:
: GENERAL INFORMATION:
:   APPLICANT: Blatner, Frederick R.
:               Burland, Valerie
:               Perna, Nicole T.
:               Plunkett, Guy
:               Welch, Rod
:
: TITLE OF INVENTION: No. 6365723el Sequences of E. coli 0157
:
: NUMBER OF SEQUENCES: 265
:
: CORRESPONDENCE ADDRESS:
:   ADDRESSEE: Charles & Brady
:   STREET: 1 South Plunkney Street
:   CITY: Madison
:   STATE: WI
:
: COUNTRY: US
: ZIP: 53701-2113
:
: COMPUTER READABLE FORM:
:   MEDIUM TYPE: Diskette, 3.50 inch, 1.44mb storage
:   COMPUTER: IBM PC compatible
:   OPERATING SYSTEM: PC-DOS/MS-DOS
:   SOFTWARE: Word Perfect 8.0

```


APPLICANT: NORRIS, STEVEN J.
 APPLICANT: JING-REN, ZHANG
 APPLICANT: HARDHAM, JOHN M.
 APPLICANT: HOMELL, JERRILYN K.
 APPLICANT: BARBOUR, ALAN G.
 APPLICANT: WEINSTOCK, GEORGE M.
 TITLE OR INVENTION: VME-LIKE SEQUENCES OF PATHOGENIC BORRELLIA
 FILE REFERENCE: US/09/125,619
 CURRENT APPLICATION NUMBER: US/09/125,619
 CURRENT FILING DATE: 1999-01-27
 NUMBER OF SEQ ID NOS: 48
 SOFTWARE: Patent In Ver. 2.1
 SEQ ID NO 3
 LENGTH: 7766
 TYPE: DNA
 ORGANISM: Borrelia burgdorferi
 FEATURE:
 NAME/KEY: modified_base
 LOCATION: (127)
 OTHER INFORMATION: R = A OR G
 US-09-125-619-3

 Alignment Scores:
 Pred. No.: 3,52e-08 Length: 7766
 Score: 186.00 Matches: 129
 Percent Similarity: 37.00% Conservative: 76
 Best Local Similarity: 23.29% Mismatches: 223
 Query Match: 7.85% Indels: 126
 Gaps: 25
 DB:
 US-09-889-314-2 (1-496) x US-09-125-619-3 (1-7766)
 QY 5 SerIleSerSerSerSerGlyProAspAsnGlnIlyAsnIleMetSerGlnValLeuThr 24
 Db 52 GCTGTAGTCTGTTAGTGGG-----GAGCAGATATTAGTCGATGTTAG 99
 QY 25 SerThrProGlnIlyValProGlnIlyAspIlySerGlyValGlnIlyThrIle 44
 Db 100 GCTGCT-----GGTGAAGCTGCCACAGATGAGACAGAGCTGGAATAATCCG 153
 QY 45 GlnGlnThrArgGlnIlyValAsnThr-----GlnMetGlu 56
 Db 154 ATTCGTGCTGCTATTGGAAGGATGATGAGAGTCCGAGTTTAAAGTGAATGAG 213
 QY 57 SerAspAlaThrIleAlaGlyAla-----SerGlyIlyAspIlyThrSer 71
 Db 214 AAGGATGATCAGATTGCTGCTATTGCTTGAAGGGGATGGCTTAAGGAGCGGACTTG 273
 QY 72 SerThrThrIlyThrGluThrAlaProGlnIlyVal-----AlaAlaGlyIlyGlu 89
 Db 274 GCGTGCAAGATGATGAGAAAGGAGGCTGAGGGGCTATTAAAGGAGCGGACTTG 333
 QY 90 SerSerGluSerGlnIlyAla-----GlyAlaAspThrGlyValSerGly 104
 Db 334 TTGGATTAAGCTGTAAACCTTAAGACACGCTGAGGGGCTTCAAGTCACTGCTGCA 393
 QY 105 AlaAlaAlaThrThrAlaSerAsnThrAlaThrIlyIleAlaMetGlnThrSerIleGlu 124
 Db 394 ATTGGAAGAGTGTGCTGATGATTAATGCTGCAAGGTTCTGATTAAGCGAGTGTAG 453
 QY 125 GluAlaSerIlySerMetGluSerThrIleGluIlySerIleu-----GlnSerIleuSer 141
 Db 454 GGGATTCTTAAGGGGATTAAGGAGATTTGAAAGCTGCGGGGAGTAAAGCTGAAA 513
 QY 142 AlaAlaGlnMetIlyGlu-----ValGlnAlaValAlaValAla 154
 Db 514 GTTGCTGCTGCTAAAGAGGCAATGAAAGCAAGGAGAGTGTGTTGGAAAGTGTAGCT 573
 QY 155 AlaIleuSerGlyIlySerSerGlySerAlaIlySerIleuGluThrProGluIleuProIlyPro 174
 Db 574 GCTCATGCTGGGACAGGAGGAGCTGCTGACAG-----606
 QY 175 GlyValThrProArgSerGluValIleGluIleGlyIleuAlaIleuAlaIlyAlaIleGln 194

Db 607 -----GGGCGTGGTCTTAGT 624
 QY 195 ThrLeu-----GlyGlnAlaThrIlySerSerAlaIleuSerAsnIlyAlaSerThrGlnAlaGln 213
 Db 625 GCTGTAGTGGGAGCAGATATTAAAGCTGATTAAGGCTGGTGGCGGCTGCTGCT 684
 QY 214 AlaAspGlnThrAsnIlySerIleuGlnIlyGln-----AlaIleIlyIleAspIly 231
 Db 685 GATCAGGAGGAGAAAGAACCTGGGATGCTAAATCCATTCGTCGCTATTGGAG 744
 QY 232 ---GluThrGlnIlyThrGlnIlyMetIlyAlaAlaGlnIlySerIlySerIlySerIly 250
 Db 745 GCTGATGCGGAGATGCTGCGGAGTTAATCATGATGAGGATGAAGAGATGAT-----798
 QY 251 GlyThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIle 270
 Db 799 -----CAGATTGCTGCTCTTATTCCTTTAGGGGAGT 831
 QY 271 SerIleValAlaAlaIle---PheThrCysGlyAlaGlyLeuAlaGlyLeuAlaIly 289
 Db 832 GCTAAGATGCAAGTTCGCTGTGAAGTGTGTGTGAGAAAGGAGGCTGAGGG 891
 QY 290 AlaAlaValAlaAlaIle-----AlaAlaGlyIlyAlaAla 301
 Db 892 GCTATTAAAGGAGCTGCTGAGTGTGATGATGCTGTAAGCTGTAAGACAGCTGAG 951
 QY 302 GlyAlaAlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAla---ValValGln 320
 Db 952 GGGGCTTCAAGTGTACATGATCATTTGAGAGATTTGGCTGCTGCTGCTGCAAG 1011
 QY 321 AlaValIlyGlnAlaValIleThrAlaValArgGlnAlaIleThrAlaIleIlyAla 340
 Db 1012 GTTGCTGATTAAGGAGCTGAGTGTGAGGAGTGTGTAAGGAGATTAAGGATGTTGAG 1071
 QY 341 AlaValIlySerGlyIlyLeuAlaPheIlyThrIleValIlyAlaIleAlaIlyAla 360
 Db 1072 GCTGGGGGGAGTAA-----AAGCTGAAGTGTGCTGCTACAGAGGAG 1116
 QY 361 IleSerIlyGlyIleSerIlyValPheAlaIly---GlyThrGlnMetIleAlaIlyAsn 379
 Db 1117 AGTATTAAGGAGGAGGAGGAGATTTGGGAGGCTGCTGCTGCTGCTGCTGCTG 1176
 QY 380 PheProIlyIleuSerIly-----ValIleSerSerIlyThrSerIlyThrValThr 396
 Db 1177 AGTGAAGCTGCTACCAAGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1236
 QY 397 ValGlyValAlaIlyValAlaIle-----404
 Db 1237 AGTGGATTT---GTTAAGGCTGCTGATGCGGCTGATCAGAGGAGAAAGCCTGGGAT 1293
 QY 405 -----AlaProAlaIleuGlyIlyGlyIlyIleMetGlnMetGlnIleuSerGlu 419
 Db 1294 GCTACAAATCCGATTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1329
 QY 420 MetGlnIlyAsnValAlaGlnPheGlnIlyGlnValGlyIlySerIleuGlnAlaAlaAsp 439
 Db 1330 AATGAGGAGATGCTGCGAGTTTAAGATGATGAGAGAGATGATGATGATGCTGCT 1389
 QY 440 MetIleSerMetPheThrGlnPheThrGlnAlaSerIlyIleAlaSerIlyGlnThr 459
 Db 1390 GCTATTGCTTTG-----AGGGGATGCTGATGATGATGATGATGATGATGAT 1443
 QY 460 GlyIlySerAsnIlyMetThrGlnIlyAlaThrIlySerIlyIlyAlaGlnIleLeu-----477
 Db 1444 GCTGAGAAAGGAGAG---GCTGAGGCGCTATTAAAGGAGCTGCTGATTTGTTGATGAT 1500
 QY 478 -----LysAlaThrAlaAlaIleSerGlyAlaIleAlaGly 489
 Db 1501 CTGGTAAACCTGTAAGACAGCTGAGGGGCTTCAAGGT 1542
 RESULT 18
 US-09-591-079-1

```
Oy 447 eTrpGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyGluSerAsnGluMetThrGln 467
Db 1634 ----GCATCAACAAGTGGCTCGGCTTCAGCAAGCACATCATGATCGATCAAGCTCAAC 1689
Oy 467 nLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyrAlaAlaIleSerGlyAlaIle 487
Db 1690 CAGTGGCTTCAGCAAGTACAGTGTCTTCAGCTTCAGCATCAACCGCGCTCGGC 1749
Oy 487 eAlaGlyAla 490
Db 1750 CTCAGCAAGC 1759

RESULT 16
US-08-961-527-365
; Sequence 365, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; COMPUTER: HP Vectra 486/33
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8512
; TELEFAX: (301) 309-8504
; INFORMATION FOR SEQ ID NO: 365:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1436 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-08-961-527-365

Alignment Scores:
Pred. No.: 1,43e-09 Length: 1436
Score: 188.00 Matches: 79
Percent Similarity: 37.30% Conservative: 59
Best Local Similarity: 21.35% Mismatches: 214
Query Match: 7.94% Indels: 18
DB: 4 Gaps: 6

US-09-889-314-2 (1-496) x US-08-961-527-365 (1-1436)
Oy 2 ThrAsnMetSerIleSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGln 21
Db 298 ACAAGTGGCTTCAGCAAGTACCTCAGCTCGAATCAGATCAACGAGTGCATCA 357
Oy 22 ValLeuThrSerThrProGlnGlyValProGlnGlnAspLysLysSerGlyAsnGluThr 41
Db 358 GCTTCAGCATCAACA-----AGTGGCTTCAGCTTCAGCAAGTATCTCAGCGCTGAATCG 411
```

```
Oy 42 LysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGluMetGluSerAspAlaThrIle 61
Db 412 GCATCAACGAGTGGCTCGGCTTCAGCAAGTACTACGGCTTCAGATCAAGTCAACCAAGT 471
Oy 62 AlaGlyAlaSerGlyLysAspLysThrSerSerThrLysThrGluThrAlaProGln 81
Db 472 GCTTCGGCTTCAGCAAGTACAGTGGCTTCAGTACAGTCAACAGTCAAGTCAAGTCA 531
Oy 82 GlnGlyValAlaIleAlaGlyLysGluSerGluSerGlnLysAlaGlyAlaLysThrGly 101
Db 532 GCAAGCATCAATCAGCTTCGATTCGATTCGATTCGATTCGATTCGATTCGATTCGATTC 591
Oy 102 ValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAla----- 114
Db 592 GCCTCAGCATCAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 651
Oy 115 -----ThrLysIleAlaMetGlnThrSerIle---GluGluAlaSerLysSerMetGlu 131
Db 652 GCGTCGACCAAGTGGCTCGGCTTCACCAAGTGCATCGATCGATCGATCGATCGATCGATCG 711
Oy 132 SerThrLeuGlnSerLeuGlnSerLeuSerAlaIleGlnMetLysGluValGluAlaVal 151
Db 712 GCCTCAGCAAGTACTAGCGCTTCAGCTCAGCATCAACAGTACAGTACAGTACAGTACAGT 771
Oy 152 ValValAlaAlaAlaLeuSerGlyLysSerSerGlySerAlaLysLeuGlnThrProGlnLeu 171
Db 772 ACTAGTCAATCAGCTTCAGCAAGTACTACAGTACAGTACAGTACAGTACAGTACAGTACAG 831
Oy 172 ProLysProGlyValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLys 191
Db 832 GCTTCAGCAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAG 885
Oy 192 AlaIleGlnThrLeuGlnGlyLysAlaThrLysSerAlaLeuSerAsnThrLysThrGln 211
Db 886 GCAAGTACCTCAGCGCTTCAGTCAAGTCAACAGTCAAGTCAAGTCAAGTCAAGTCAAGTCA 945
Oy 212 AlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGln-----LysGlnAlaIleLys 228
Db 946 GCATCAAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTAC 1005
Oy 229 IleAspLysGluArgGlnGluGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 248
Db 1006 GCATCAACGAGTGGCTTCAGTTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAGGCTCAG 1065
Oy 249 LeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThr 268
Db 1066 GCCTCGGCTTCAGCAAGTACAGTGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAG 1125
Oy 269 ValIleSerIleValAlaIleAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAla 288
Db 1126 GCATCAACGAGTGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAG 1185
Oy 289 GlyAlaIleValGlyAlaAlaIleAlaIleGlyAlaIleGlyAlaIleAlaIleAlaIleThr 308
Db 1186 GCGTCAGCTTCGCGATCAACAGCGCTCGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCTTC 1245
Oy 309 ValAlaIleThrGlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaIleThr 328
Db 1246 GCATCAACGAGTGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAGGCTTCAG 1305
Oy 329 AlaValArgGlnAlaIleThrAlaIleAlaIleLysAlaAlaValLysSerGlyIleLysAla 348
Db 1306 GCGTCGGCTTCAGCAAGTACCAAGTGGC---TCAGCTTCAGCAAGTACCAAGTGGTCAAGCT 1362
Oy 349 PheIleLysThrLeuValLysAlaIleAla 358
Db 1363 TCAGCATCAACCAAGTGGCTTCGCGCTTCGCGCA 1392

RESULT 17
US-09-125-619-3
; Sequence 3, Application US/09125619
; Patent No. 6437116
; GENERAL INFORMATION:
```

```

US-08-961-527-364
; Sequence 364, Application US/08961527
; Patent No. 6420135
; GENERAL INFORMATION:
; APPLICANT: Charles Kunsch
; TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
; NUMBER OF SEQUENCES: 391
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Human Genome Sciences, Inc.
; STREET: 9410 Key West Avenue
; CITY: Rockville
; STATE: Maryland
; COUNTRY: USA
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
; OPERATING SYSTEM: MSDOS version 6.2
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/961,527
; FILING DATE:
; CLASSIFICATION: 424
; PRIOR APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Brookes, A. Anders
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PB340P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (301) 309-8504
; TELEFAX: (301) 309-8512
; INFORMATION FOR SEQ ID NO: 364:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2550 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
;
US-08-961-527-364
;
Alignment Scores:
Pred. No.: 1.43e-09 Length: 2550
Score: 192.00 Matches: 110
Percent Similarity: 33.97% Conservative: 68
Best Local Similarity: 20.99% Mismatches: 275
Query Match: 8.10% Indels: 71
DB: 4 Gaps: 10

US-09-889-314-2 (1-496) x US-08-961-527-364 (1-2550)
QY 5 SerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeuThr 24
DB 289 TCACACATCGCTCGCTCGATCGACATCGACAGCGCTCGCTCGACGACGATCCAGTACCGATGCT 348
QY 25 SerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGlnIle 44
DB 349 TCAGCTCGACGCTCGACAGATCGCTCGGCTTCACACCATGATCGATCGATCGATCGATCAACC 408
QY 45 GlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGlyAla 64
DB 409 AGTGGCTCGACCTCGACAGTACTAGTCATCGATCGATCGATCGATCGATCGATCGATCGGCT 468
QY 65 SerGlyLysAspLysThrSerThrThrThrLysThrGlnThrAlaProGlnGlnGlyVal 84
DB 469 TCAGCATCAACACAGTCCCTCGCTCGATCGATCGATCGATCGATCGATCGATCGATCGATCGGCT 528
QY 85 AlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThrGlyValSerGly 104
DB 529 AGTGGCTTCAGTTCAGCATTAACAGTGGCTCGATCGATCGATCGATCGATCGATCGATCGGCT 588
QY 105 AlaAlaAlaThrThrAlaSerAsnThrAla---ThrLysIleAlaMetClnThrSerIle 123

```

```

DB 589 TCAGCAAGCACATTCAGATCTGATCAGCTCGACAGATGCGCTCGGCTCAACACAGTCA 648
QY 124 ---GlnGlnLysSerSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerIle 142
DB 649 TCCTAATCGGCATCAACCATGCTCGCTCGATCGATCGATCGATCGATCGATCGATCGATCGGCT 708
QY 143 AlaGlnMetLysGlnValGlnValAlaValAlaAlaAlaSerGlyLysSerGly 162
DB 709 TCACGAGTGCATGCGCTCGCTCGATCAACAGTGCATCGATCGATCGATCGATCGATCGGCT 768
QY 163 SerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnVal 182
DB 769 TCAGCTTCGGCATCAACAGT----- 789
QY 183 IleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnLysAlaThrLysSer 202
DB 790 -----GCTCGCTTCAGCAAGCACATCACTGATCTGATCAAGCTCGGCTCAAC 834
QY 203 AlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeu 222
DB 835 AGTGGCTCGCTTCAGCAAGTACAGTCTCGATCGATCGATCGATCGATCGATCGGCTCAAC 888
QY 223 GlnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGln 948
DB 889 TCAGCTCGACAGCACCTCAGCTTCTGATCGCTCGATCGATCGATCGATCGATCGGCTCGA 948
QY 243 GlnGlnLysSerLysAspLeuGlnLysThrMetAspThrValAsnThrValMetIleAla 262
DB 949 AGCACTCAGCTTCGATCGCTCGATCGATCGATCGATCGATCGATCGATCGATCGATCGGCT 1008
QY 263 ValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyValGly 282
DB 1009 TCAGCTTCAGCAAGCACAGCGCTCGCTCGATCGATCGATCGATCGATCGATCGGCTCGA 1068
QY 283 LeuAlaGlyLeuAlaLysAlaLysAlaLysAlaLysAlaLysAlaLysAlaLysAla 297
DB 1069 TCACACATGCTCGCTCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGGCT 1128
QY 298 GlyLysAlaAlaLysAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 317
DB 1129 TCTGATCGGCATCAACGATCGCTCGATCGATCGATCGATCGATCGATCGATCGATCGGCT 1188
QY 318 ValValGlnAlaValLysGlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAlaAla 337
DB 1189 AGCACATGCTTCAGCTCGCTCGATCGATCGATCGATCGATCGATCGATCGATCGATCGGCT 1248
QY 338 IleLysAlaAlaValLysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIle 357
DB 1249 ---TCGCTTCAGCAAGTACTAGCGCTCGATCGATCGATCGATCGATCGATCGATCGGCT 1305
QY 358 AlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAla 377
DB 1306 GCG-----TCACAGATGCTGCTCGATCGATCGATCGATCGATCGATCGATCGATCGGCT 1356
QY 378 LysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysThrValThrVal 397
DB 1357 AGCACATGCTTCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGGCT 1416
QY 398 GlyValGlyValValAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 417
DB 1417 TCAGCTTCAGCAAGTACAGTCTCGATCGATCGATCGATCGATCGATCGATCGATCGGCT 476
QY 417 userGlnMetGln-----GlnAsnValAlaGlnPhe 427
DB 1477 TCACACAAGTGCCTCGCTCGATCAACAGTGCATCTGATCGATCGATCGATCGATCGGCT 429
QY 427 eGlnLys----- 429
DB 1537 CAGCAAGTACTAGTGCATCGATCGATCGATCGATCGATCGATCGATCGATCGATCGGCT 1596
QY 430 -----GlnValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPhe 447
DB 1597 GTGCATCAGAGTCAAGCAAGTACCAAGTGCATCTTC----- 1633

```

NUMBER OF SEQUENCES: 391
 CORRESPONDENCE ADDRESS:
 ADDRESSER: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/961,527
 FILING DATE:
 CLASSIFICATION: 424
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Brookes, A. Anders
 REGISTRATION NUMBER: 36,373
 REFERENCE/DOCKET NUMBER: PB340P1
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (301) 309-8504
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ ID NO: 367:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1702 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-961-527-367
 Alignment Scores:
 Pred. No.: 5, 87e-10 Length: 1702
 Score: 193.00 Matches: 112
 Percent Similarity: 37.85% Conservative: 75
 Best Local Similarity: 22.67% Mismatches: 267
 Query Match: 8.15% Indels: 42
 Gaps: 9
 US-09-889-314-2 (1-496) x US-08-961-527-367 (1-1702)
 QY 19 MetserglValleuthrserThrProglngllyValProglnglnasplysleuSerGly 38
 Db 279 CTTCAGCAAGTACAGTGCCTCAGCTCAGCAAGTACAGTGCCTCAG----- 326
 QY 39 AsnluThrlysglnlleglnlThrArglnlGlylysAsnThrglumetGlySerasp 58
 Db 327 -----CCTCAGCTGCACCAAGTGCCTGCGCC-TCACCACTGCATCTGAATCGGCA 376
 QY 59 AlathrllealaglylaserglyLysAsplysThrserSerThrThrluThr 78
 Db 377 TCACCAAGTGGGTGAGCTGCACCAAGTACTGTCATCAGCTTCAGCATCAACAGAGCA 436
 QY 79 AlAProglngllyValAlaAlaGlyLysGlySerSerGlySerGlyLysAlaGlyAla 98
 Db 437 TCGGCTTCAGCATCAACAGTGCATCAGAGTCAAGTACAGTGCCTGCTCGGCA 496
 QY 99 AsphrglyValserGlyAlaAlaAlaThrThrlaserAsnThrla----- 114
 Db 497 TCACAAGTGGCTCGCTTCAGCAAGTACTGAGCGCTCAGCTCAGCTCAACAAGTCT 556
 QY 115 -----ThrlsIlleAlaMetGlnThrSerle-----GluGlnAla 126
 Db 557 TCAGCTTCGGGTGACAGAGCGCTCGGCTTCACCAAGTATCTCAGCTCGATTCGGCA 616
 QY 127 SerlyserMetGlySerThrlleuglnSerleuglnSerleuSerAlaAlaGlnMetLys 146
 Db 617 TCACAAGTGGCTCGCTTCAGCATCAACGAGTGCATCAGTCTCAGCAACAGCAGTGGC 676

QY 147 GluValGluAlaValAlaAlaAlaAlaSerGlyLysSerSerGlySerAlaLysleu 166
 Db 677 TCGGCTTCAGCAAGTACAGAGCGGCTGTAATCGCATCAACAGTGCCTCAGCTTACGA 726
 QY 167 GluThrProgluLeuProlyserProgllyValThrProArgSerGluValleuglnGly 186
 Db 737 AGTACCTCAGCATCTGATCAGCATCAACAGTGCATCGGCTTCAGCAACAGCA----- 790
 QY 187 LeuAlaLeuAlaLysAlaAlleGlnThrlleuglnGluAlaThrlsSerAlaLeuSerAsn 206
 Db 791 AGTGCCTCAGCTTCAGCAAGTATCTCAGCTGTAATCGCATCAACAGTGCCTCAGCTTACGA 850
 QY 207 TyrAlaSerThrglnAlaGlnAlaAspGlnThrasnLys-LeuGlyLeuGlnLysGlnAl 226
 Db 851 TCAGCAAGTACTAGCGCTCAGCATCAGCTCAACAAGTCTTCGGCTTCAGGCTCAACG 910
 QY 226 alleLysIlleAsplysGlu-----ArgGluGluTyrGlnGlnMetLysAlaAlaGly 243
 Db 911 AGTGCCTCAGTACAGCATCAACAGTGCATCGGCTTCAGCAAGC-ACATCAGCTTCTGA 969
 QY 243 uGlnLysSerLysAspLeuGlnThrlMetAspThrlValAsnThrValMetlleAlaVala 263
 Db 970 ATCTGCATCAACCAAGTGCCTCAGCTCAGCATCAGCAAGCGCTTCAGCATCAGCATC 1029
 QY 263 lSerValAlaAlleThrVallleSerlleValAlaAlaAllePheThrcysGlyAlaGlyLe 283
 Db 1030 CAGTGCCTCA-----GCCCTCAGCAAGTACCAAGTGCCTCAGCTC 1068
 QY 283 uAlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaAlaGlyAlaAlaGlyAla 303
 Db 1069 AGCGTTCAGCAAGTGCCTGCGCTTCACCAAGTGCATGTAATCGCATCAACAGTGCCTC 1128
 QY 303 aAlaAlaAlaThrThrValAlaThrlGlnlleThrValGlnAlaValAlaGlnAlaVal 323
 Db 1129 AGCTTCAGCAAGTACTAGCGCC-----TCAGCTCAGCATCAACAGTGCCTCAGCTC 1179
 QY 323 sGlnAlaVallleThrAlaValAlaArgGlnAlaAlleThrAlaAlaAlleLysAlaAlaVal 343
 Db 1180 TTCAGCAAGTACTAGTGCATCAGCATCAGCATCAGCATCAACAGAGCA--TCGGCTTCAGCAAG 1226
 QY 343 sSerGlylleLysAlaPheIleLysThrlleuValLysAlaAlleAlaLysAlaAlleSerly 363
 Db 1237 TACAGCGCTCAGCTTCAGCAAGTGCCTCAGCTCAGCTCAGCTCAGCAAGTACAGCGCTC 1296
 QY 363 sGlylleSerLysValPheAlaLysGlyThrlGlnMetlleAlaLysAsnPhelProlyle 383
 Db 1297 AGCTTCAGCAAGCAACAGTGCCTCAGCTCA-----GCAAGTACCAAGTGCCTCAGC 1347
 QY 383 userLyVallleSerSerleuThrSerLysTrpValThrValGlyValAlaVala 403
 Db 1348 CTCAGCTTCAGCAAGTGCCTGCGCTTCAGCAAGTACTCAGCTCAGCTTCGAATCAGCATCAAC 1407
 QY 403 lAlaAlaProAlaLeuGlyLysGlylleMetGlnMetGlnLeuSerGlnMetGlnGlnas 423
 Db 1408 GAGTGCATCAGCTTCAGCATCAACAGTGCCTCAGCTTCAGCAAGTGCCTGCGCTCAGCTCAG 1467
 QY 423 nValAlaGlnPheGlnLysGluValGlyLysLeuGlnAlaAlaAlaAlaAlaAlaAlaAla 443
 Db 1468 TTCAGCATCAACAGTGCCTTCAGCTCAGCTCAGCTCAACAGTGCCTTCGAATCAGCATCAAC 1527
 QY 443 tPheThrlGlnPheTrpGlnAlaSerLyslleAlaSerLysGlnThrlGlyGlnSerAs 463
 Db 1528 AAGTGCCTCGGCTTCAGCAAGTGCCTGCGCTTCAGCAAGTGCCTTCGAATCAGCATCAAC 1587
 QY 463 nGlnMetThrlGlnLysAlaThrlLysLeuGlyAlaGlnlleuLysAlaTrpAlaAlaAl 483
 Db 1588 TTCAGCATCAACAGTGCCTTCGAATCGGCTTCAGCAAGTGCCTTCGCTTCAGCATCAAC 1647
 QY 483 eSerGlyAlaAlleAlaGly--AlaThrlsThrasnAsn 495
 Db 1648 GAGTGCCTCAGCTTCAGCAAGTGCATCAGTCTTGAAT 1685
 RESULT 15

APPLICATION NUMBER: 60/110,955
FILING DATE: 04-DEC-1998
ATTORNEY/AGENT INFORMATION:
NAME: Seay, Nicholas J.
REGISTRATION NUMBER: 27386
REFERENCE/DOCKET NUMBER: 960296.95017
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 251-5000
TELEFAX: (608) 251-9166
INFORMATION FOR SEQ ID NO: 72:
SEQUENCE CHARACTERISTICS:
LENGTH: 46819
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
SEQUENCE DESCRIPTION: SEQ ID NO: 72:
US-09-453-702B-72

Alignment Scores:

Pred. No.:	1,23e-08	Length:	46819
Score:	203.00	Matches:	127
Percent Similarity:	37.33%	Conservative:	66
Best Local Similarity:	24.56%	Mismatches:	208
Query Match:	8.57%	Indels:	116
DB:	4	Gaps:	20

US-09-889-314-2 (1-496) x US-09-453-702B-72 (1-46819)

QY 5 SerIleSerSerSerSerGlyProAspAsnGlnIlyAsnIleMetSerGlnIleValLeuThr 24
DB 32174 TCAGGCTCTTCACGCGCAGGA-----ACGGCA 32200
QY 25 SerThrProGlnIlyValProGlnIlyAsnIlySerGlyAsnGlnIlyThrIle 44
DB 32201 TCACAAAGGCTACTGAGCATCAAAAGTGGCTGCCGTCAGAGTCTCCCAAAAGCGCG 32260
QY 45 GlnGlnThrArgGlnIlyLysAsnThrGlnMetGlnUserAspAlaThrIleAla 64
DB 32261 GCGGCTACCAAGCGCGGTCGCGGCAAAACGTCAGAAACGCAATGCGCACTGTCCACACA 32320
QY 65 SerGlyLysAspIlySerSerSerSerThrIlyThrGlnIlyThrAlaProGln----- 81
DB 32321 TCAGCGCGCATTCGTGATCCACCGCAGCAAGCGTCAGAGTCTCCCTCAGCC 32380
QY 82 GlnGlnIlyValAlaIlyGlnIlySerIlySerGlnIlySerGlnIlyAlaIlyAspThrIly 101
DB 32381 AGGATGCGCTCGCTCAAAAGAGCGCGCAAAATCATCAGAAACGACGCGCCTCCGAC 32440
QY 102 ValSerGlyAlaAla-----AlaThrThrAlaSerAsnThrAlaThrIlyIleAlaMet 119
DB 32441 GCCATAGTCAGGCTCTCCGCGCAACGCGCGCAATTCGCGCAAG-----GCGGCC 32494
QY 120 GlnThrSerIleGlnIlyLysIlySerIlySerMetGlnIlySerThrLeuGlnIlySer 139
DB 32495 AAAAGCTCT---GAGACAAACCTTAAGTCTCTGAAACGCGACAGAA-----CAGAGT 32545
QY 140 LeuSerAlaIlyAlaIlyMetIlySerIlyGlnIlyAlaIlyValIlyAlaIlyLeuSerIly 159
DB 32546 GCGCTCCGCGACGAGCGCTCAAAACAGCGCGCTGATATCTGCGCCTCCGCTCAACA 32605
QY 160 SerSerGlySerAlaIlyLeuGlnIlyThrProGlnIlyProIlyProGlnIlyValThrProArg 179
DB 32606 AGTGGCGGCGAGCC-----TCAGCCAGTGGCAGCGCGCGCGGAAATTCGCGCA 32656
QY 180 SerGlnIlyAlaIlyGlnIlyLeuAlaIlyAlaIlyAlaIlyGlnIlyThrLeuGlnIlyAla 199
DB 32657 AGTGGCGGCTCTGCTCTCAACAGCACAACGAGGCT-----GGCGAGCC 32704
QY 200 ThrIlySerAlaIlyLeuSerAsnThrIlyAlaSerThrGlnAlaIlyAlaIlyAspIlyThrAsn 219
DB 32705 ACTGAACAGCGCAGCGACGACGAGTCTGCTCCGCA----- 32743

QY 220 LeuGlyLeuGlnIlyGlnAlaIlyIleLysIleAspIlySerGlnIlyGlnIlyThrIle 239
DB 32744 -----GCG 32746
QY 240 LysAlaIlyGlnIlyLysSerIlyAspLeuGlnIlyThrMetAspThrValAsnThrVal 259
DB 32747 AAGACATCCGAACGAAACGAAACGCTTGGAACACGCGCAAGATCTCTCAAAAACG--- 32803
QY 260 MetIleAlaIlySerValAlaIlyIleThrValIlySerIleValAlaIlyIleThrIly 279
DB 32804 -----GCTGCCGCTCTGCTCAACCCT 32824
QY 280 GlyAlaIlyLeuAlaIlyLeuAlaIlyAlaIlyAlaIlyAlaIlyAlaIlyGly 299
DB 32825 TCGCGCGGCTCATCGCATCTCTGCTCTCAAAAAGATGAGCGCAGCAACGACG 32884
QY 300 AlaAlaIlyAlaIlyAlaIlyAlaIlyThrValAlaIlyThrGlnIlyIleThrValIlyAlaIly 319
DB 32885 TCAGCAGCGAAGACGACGCGCAGCAGCATTCACGAGCGCAGCGCAGCGCT----- 32938
QY 320 GlnAlaIlyLysGlnAlaIlyIleThrAlaIlyAlaIlyAlaIlyIleThrAlaIlyIly 339
DB 32939 -----GATGTCGCGACGCGCAGCGCTCAGACGCAAAAGTACGCGGAAATCT 32983
QY 340 AlaAlaIlyLysSerGlyIlyLeuAlaIlyIleLysIleThrValIlyAlaIlyAlaIly 359
DB 32984 GCAGCA-----ACGCGCGCTGACACAGCGGCAAAA 33013
QY 360 AlaIleSerLysGlyIlySerLysValPheAlaIlyGlnIlyGlnIlyMetIleAlaIlyAsn 379
DB 33014 CCGCGAGAGATATTCATTCGCGCGCTGAGATGCGACGACGACGCAAAAGGG 33073
QY 380 PheProIlyLysSerLysValIleSerSerIleThrIlySerIlyThrValIlyAlaIly 399
DB 33074 ATGATACAGCTCAGCAGTGCAGCTAAGC---ACTTCGAGTCTGCGCGCAAGCCCA 33130
QY 400 GlyValIlyAlaIlyAlaIlyProAlaIlyLeuGly----- 409
DB 33131 AAAGCGTTAAGCGCGCTGATGCGCTAAGCGGAAATACCGCAGCATGCGACG 33190
QY 410 -----LysGlyIlyMetGlnMet-----GlnLeuSerGlnIlyMetGln 421
DB 33191 ACAGCAGAAAGGATGATGCTAGCAACGCGACACGACGACGATCTGAAATG--- 33247
QY 422 GlnAsnValAlaIlyGlnIlyPheGlnIlyGlnIlyGlnIlyGlnIlyAlaIlyAlaIly 441
DB 33248 -----CTGGCGGCAACGCCCAAGTCGGA-----AAGCGACCTATGACCTTCT 33292
QY 442 SerMetPheThrGlnPheThrGlnIlyAlaIlySerIlyIleAlaIlySerGlnIlyGlnIly 461
DB 33293 AAC-----GGAAATATATCTCTCAGACGCTACGACAGCACAACAAAGAAATTCAG 33346
QY 462 SerAsnGlnMetThrGlnIlyAlaIlyThrIlySerIlyGlnIlyAlaIlyIleLeuIlyAlaIly 481
DB 33347 CTCACATAGTCAACCAACAGCAGCATCTGAACGCTTGGCGGACACCAAGCACTGTAA 33406
QY 482 Ala-----IleSerGlyAlaIlyAlaIlyGlnIlyAlaIlyHisIlyThrAsn 494
DB 33407 CGAGCTAATGATATGATGCAATGTCGGGTACCTTCTGCTTGAAGTGAAT 33457

RESULT 13
US-08-961-527-363
Sequence 363: Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland

Db 936 GCTCTGGA-----GAGGCTCAGAGGCGAGCTCTCTATGAGCCAGCATCAA 986
Qy 68 AsplvsthrserSerthrrhrthrlglnthrlaProglnglnlyValaAlaAgly 87
Db 987 -----AAGACGATACCCGCAAGAGTGTATGACCTGCGGCC 1025
Qy 88 LysgluserSerluserGlnlySala-----GlyAlaAspThrGlyVal 102
Db 1026 AAAAACTGACGAGCGGCAATAATTCACATCGCTGACCCAGCGCTAT 1085
Qy 103 SerGlyAlaAlaAlaThrthrlaSerAsnThrAlaThrLysLleAlaMetLthrSer 122
Db 1086 GCACAGCTGAGCC-----GCG 1103
Qy 123 lIeGlunAlaSerlySerMetGluSerThrlenglnSerluserSera 142
Db 1104 GTAGAACAGCCGCAAGAACGAGCGCAAGAGCCCTTAGATAGCCAGCAT 1163
Qy 143 AlaGlnMetLys-----GluValGlnAlaValaValaAlaAlaLuser 157
Db 1164 GCGAGGTTAAAGCAGCAGCAGCCCAAGCAAGCAAGCGATAACATTCTG 1223
Qy 158 GlyLysSerSerGlySerAlaLysLenglnThrProgluLeuProLysProGlyValThr 177
Db 1224 ACCAATTCAGGGAACGCTAATCGCGCTCAGATCAGTTCCAGGCTGAGCAG 1283
Qy 178 ProArg-----SerGluValIleGlnIleGlnAlaLysAlaAlaIleGlnThr 196
Db 1284 GATATCTGTCAATGTCCCGCCCTCACTATGCTCATGCGCATGTTATGAGATTGTC 1343
Qy 197 GlyGlnAlaThrLysSerAlaLuserAenryrAlaSerThrglnAlaAlaSpLn 216
Db 1344 GCGAAATATAGGAGAAACCTGCAAAACGATCTTGCCCT-----1385
Qy 217 ThrAsnLysLenglnLysGlnAlaIleLysIleAspLysGlnArgGlnGlnTyr 236
Db 1386 TTCACCCCTTCGCAAGAGCGCTCAGCGC---GAGATGAAAGAAAGACGCTGAATTC 1442
Qy 237 GlnGlnMetLysAlaAlaGlnLysSerLysAspLenglnGlyThrMetAspThrVal 256
Db 1443 CAGGAA-----GAGACCGCAAGCCGAGAGAAACGACGATATGAGATGATC 1493
Qy 257 AsnThrValMetIleAlaValaSerValAlaIleThrValIleSerIleValAlaAlaIle 276
Db 1494 GGGAAAGTCCTCGCGCGCTG-----CTAACATTCGTGACGCTTGCGCGCTGTT 1544
Qy 277 PheThrCysGlyAlaGlyLeuAlaGlyLeuAlaGlyAlaAlaValaGlyAlaAlaAla 296
Db 1545 TTTACCGGTGGGCGAGTCTGCGCTGCGTGGCGACTTGGCGTA-----1592
Qy 297 AlaGlyGlyAlaAlaGlyAlaAlaAlaAlaThrValAlaThrGlnIleThrValGln 316
Db 1593 -----ATGTCGCCGATGAAT-----1610
Qy 317 AlaValaGlnAlaValaLysGlnAlaVal-----IleThrAlaValaArgGlnAlaIleThr 335
Db 1611 -----GTGAAGCGCGCGAGCGGGGTGCTGTTATTCACAGCCCTAAMC 1655
Qy 336 AlaAlaIleLysAlaAlaValaLysSerGlyIleLysAlaPheIleLysThrLeuValLys 355
Db 1656 CCGATTATGAGCATG-----CTGAAGCCGTATATGAG 1691
Qy 356 AlaIleAlaLysAlaIleSerLysGlyIleSerLysValaPheAlaLysGlyThrGlnMet 375
Db 1692 CTGATTGGCAGCGGATTCACAAAGCGCTG-----GAAGCATTAAGCGCTC 1736
Qy 376 IleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysTrpVal 395
Db 1737 CATAGAAACGCGAGATGCGAGCGCATTTGTCGATGTCGCCGTATGTC 1796
Qy 396 ThrValGlyValaGlyValaVal-----AlaAlaProAlaLeuGly 409
Db 1797 ATGTAGCGCTCATTTGTGTGTCGACATTGCGGAAGCGCGCGCAATGCGGT 1856

Qy 410 LysGlyIleMetGlnMetGlnLeuSerLysMetGlnAsnValaAlaGlnPheGlnLys 429
Db 1857 AACGGCTGACCAAAAT-----ATGGCGCAACGATTAGAGATTGGGCT 1904
Qy 430 GluValGlyLysGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPheTrpGln 449
Db 1905 AAGTGCCTGAAA---CACTGGCACAAACGAGCAAACTCTTACCCAGGGATGCAA 1961
Qy 450 Gln-----AlaSerLysIleAlaSerLysGlnThrGlyGln 461
Db 1962 CGATTACTAGCGCGCTGCTGTAATGTGGTAGCAAGATGGCTGCAAGCATGCTTA 2021
Qy 462 SerAsnGlnMet-----ThrGlnLysAlaThrLysLenglnAlaGlnIleLeuLys 478
Db 2022 AGTAAGAGCTGTGATATACCTAAATAAGTGGCGTTGGCAGTAGAGTCAGCAAT 2081
Qy 479 AlaThrAlaAlaIleSerGlyAlaAlaAla 488
Db 2082 ACCGACGCCAGTCAGCGCGGTGGCTGCC 2111
RESULT 10
US-08-591-079-7
Sequence 7, Application US/08591079
Patent No. 5972899
GENERAL INFORMATION:
APPLICANT: Zychlinsky, Arturo
APPLICANT: Chen, Xajing
TITLE OF INVENTION: Apoptosis Induced by Shigella Ipab
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,079
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Livanac, Samuel 33,949
REGISTRATION NUMBER: 15661-20017.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0764
TELEX: 90-4030 MRSNFOERSMWH
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3622 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
ORIGINAL SOURCE:
ORGANISM: Salmonella typhimurium
STRAIN: slbB
FEATURE:
NAME/KEY: CDS
LOCATION: 575..2356
OTHER INFORMATION: /product= "slpB"
OTHER INFORMATION: /gene= "slpB"
US-08-591-079-7
Alignment Scores: 5,26e-12 Length: 3622
Pred. No.: 218.50 Matches: 116
Score:

APPLICATION NUMBER: JP 106011/95
FILING DATE: 28-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Miller, Charles E.
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 7426-043-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 18:
SEQUENCE CHARACTERISTICS:
LENGTH: 1296 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid; Synthetic DNA
US-08-809-326A-18

Alignment Scores:
Pred. No.: 1,44e-116 Length: 1296
Score: 1245.00 Matches: 259
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 52.55% Indels: 0
DB: 4 Gaps: 0

US-09-889-314-2 (1-496) x US-08-809-326A-18 (1-1296)

QY 4 MetSerIleSerSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23
DB 484 ATGTCTATTTCATCTCTCTCAGAGACTGACAAATCAAAAAATATCATGTCTCAAGTCTCG 543
QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyValnGlnThrLysGln 43
DB 544 ACATGCACACCCAGGCGCTGCCCAAGATTAACCTGTCTGGCAACGAAGCAAGCA 603
QY 44 IleGlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63
DB 604 ATACAGCAACACGCTCAGGCTAAAAACACTGACATGGAAGCATGCCATCTATTGCTGGT 663
QY 64 AlaSerGlyLysAspLysThrSerSerThrLysThrLysGlnThrAlaProGlnGlnGly 83
DB 664 GCTTGGGAAAGACAAACTCTCTGACTACAAAAACAGAAACAGCTCCACAAAGGGA 723
QY 84 ValAlaAlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThrGlyValSer 103
DB 724 GTTGCTGTGGGAAAGAAATCTCAGAAAGTCAAAAGGAGTCTGATCTGAGATATCA 783
QY 104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
DB 784 GGAAGCGGCTGCTACTACACATCAATACCTGCAACAAAATGTCTATGACAGACTCTATT 843
QY 124 GlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143
DB 844 GAAGAGCGAGCAAAAGTATGAGTCTACCTTAGACTCACTCAAGCCCTCAGTCCGGG 903
QY 144 GlnMetLysGlnValGlnAlaValValAlaAlaLeuSerGlyLysSerSerGlySer 163
DB 904 CAATGTAAAGAGTCGACACGGTGTGTGCTGCCCTCAGGAAAGATTGGGTTCC 963
QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnValIle 183
DB 964 GCATAATTTGGAACCTTAGCTCCCAAGCCCGGGTGCACCAAGATCAGAGTTATC 1023
QY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnAlaThrLysSerAla 203
DB 1024 GAAATCGAGCTGCGCTCTAAAGCAATTCACACATTTGGAGAACCCCAAAATCTGCC 1083
QY 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGln 223
DB 1084 TTATCTAATGATGACAGTACACAAAGCAAGACGACCAAAACAATTAACGTATGAGAA 1143

QY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlyThrGlnGlnMetLysAlaAlaGln 243
DB 1144 AAGCAAGCATATAAATGATTAAGACAGAGAAAGATACCAAGAGATGAAGCTGCCGA 1203
QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAla 262
DB 1204 CAGAAGCTTAAGATCTCGAAGCAACATGATCTCAATACTGTGATGATCGCG 1260

RESULT 9

US-08-591-079-9
Sequence 9, Application US/08591079
Patent No. 5972899
GENERAL INFORMATION:
APPLICANT: Zychlinsky, Arturo
TITLE OF INVENTION: Apoptosis Induced by Shigella IPAB
NUMBER OF SEQUENCES: 10
CORRESPONDENCE ADDRESS:
ADDRESSEE: MORRISON & FOERSTER
STREET: 2000 Pennsylvania Avenue, NW
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20006-1812
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/591,079
FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:
NAME: Litvat, Shmuel
REGISTRATION NUMBER: 33,949
REFERENCE/DOCKET NUMBER: 15661-20017.00
TELECOMMUNICATION INFORMATION:
TELEPHONE: (202) 887-1500
TELEFAX: (202) 887-0764
TELEX: 90-4030 MRSNFOERSWSH

INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 5393 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: DNA (genomic)
ORIGINAL SOURCE:
ORGANISM: Salmonella typhi

FEATURE:

NAME/KEY: CDS
LOCATION: 543..2324
OTHER INFORMATION: /gene="sfpB"
US-08-591-079-9

Alignment Scores:
Pred. No.: 6.3e-12 Length: 5393
Score: 220.50 Matches: 116
Percent Similarity: 41.02% Conservative: 85
Best Local Similarity: 23.67% Mismatches: 182
Query Match: 9.31% Indels: 107
DB: 2 Gaps: 20

US-09-889-314-2 (1-496) x US-08-591-079-9 (1-5393)

QY 28 GlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGlnIleGlnGlnThr 47
DB 876 CAGGCAATGTTGATGACAAAGAGATGGGATTCAGATCGAAGCAATTCAGACG 935
QY 48 ArgGlnLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGlyAlaSerGlyLys 67

STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,326A
FILING DATE: 19-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 224711/94
FILING DATE: 20-SEP-1994
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106006/95
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106008/95
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106009/95
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106010/95
FILING DATE: 28-APR-1995
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 106011/95
FILING DATE: 28-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Miller, Charles E.
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 7426-043-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEO ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 777 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Genomic DNA
US-08-809-326A-7
Alignment Scores:
Pred. No.: 6.3e-117 Length: 777
Score: 1245.00 Matches: 259
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 52.55% Indels: 0
DB: 4 Gaps: 0
US-09-889-314-2 (1-496) x US-08-809-326A-7 (1-777)
QY 4 MetSerIleSerSerSerSerGlyProAspAsnGlnIlyAsnIleMetSerGlnValIleu 23
DB 1 ATGCTATTTCATCTTCTTCAGGACCTGACATCAAAAAATATATCATGCTCAAGTTCTG 60
QY 24 ThrSerThrProGlnGlnValIleProGlnIleAspLysLeuSerGlyAsnGlnIleThrLysGln 43
DB 61 ACATCGACACCCCGGCGGCGCCCAACAGATAGCTGTCTGCAACGAAACGAAACGAA 120
QY 44 IleGlnIleThrArgGlnGlyLysAsnThrGluMetGlnSerAspAlaThrIleAlaGly 63
DB 121 ATACAGCAACACGCTACGAGTAAACACATGAGATGAAAGCCATGACATATGCTGCT 180
QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly 83
DB 181 GCTTGTGAAAAAGCAAAACTCTCTGACTACAAAAACAGAAACAGTCCACACAGGGA 240

QY 84 ValAlaIleGlyLysGlnSerSerGlnSerGlnLysAlaIleAspThrGlyValSer 103
DB 241 GTTGCTGCTCGGAAAGATCTCTCAGAAAGTCAAAGCCAGCTGCTATCTGAGATATCA 300
QY 104 GlnValAlaIleThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
DB 301 GGAGCGGCTGCTACTACAGCATCAATATCGCAACAAATATGCTATGACAGCTCTATT 360
QY 124 GlnGlnAlaSerLysSerMetGlnSerThrLengGlnSerLeuGlnSerSerAlaAla 143
DB 361 GAAGAGCGGAGCAAAAGATATGAGTCTACTTACCTTAGAGTCACTTCAAAAGCTCAGTCCGC 420
QY 144 GlnMetLysGlnValGlnAlaValAlaIleAlaIleAlaSerGlyLysSerSerGlySer 163
DB 421 CAATGAAAGATCGAAGCGGTTGTTGTTGCTCCCTTCAGAGGAAATTCGGGTTCC 480
QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnValIle 183
DB 481 GCAAAATTTGGAAACACCTGAGCTCCCAAGCCGGGGTGACACCAAGATCAGAGTTATC 540
QY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLengGlyGlnAlaThrLysSerAla 203
DB 541 GAAATCGAGCTCGGCTTGTAAAGCAATTCAGACATTTGGAGAGACCAAAATTCGCC 600
QY 204 LeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGln 223
DB 601 TTATCTACTATATGCAAGTACCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 660
QY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnIleThrLengLysAlaIleGln 243
DB 661 AAGCAAGCGATAAATATCGATTAAGAGAGAGAAATACCAAGATGAGAGCTGCGCGAA 720
QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAla 262
DB 721 CAGAGTCTAAAGATCTCGAAGAGCAATGATGCTGATCTGATGATCGCG 777
RESULT 7
US-08-809-326A-4
Sequence 4' Application US/08809326A
Patent No. 6165478
GENERAL INFORMATION:
APPLICANT: Izutsu, Hiroshi
APPLICANT: Obara, Kazuhiko
APPLICANT: Matsunoto, Akira
TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES,
TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING
TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMA
TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION O
TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREM
TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA
TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND
TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND
TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
TITLE OF INVENTION: PNEUMONIAE GENE
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,326A
FILING DATE: 19-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 224711/94

CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08-809,326A
 FILING DATE: 19-MAR-1997
 CLASSIFICATION: 435
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 224711/94
 FILING DATE: 20-SEP-1994
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 106006/95
 FILING DATE: 28-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 106008/95
 FILING DATE: 28-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 106009/95
 FILING DATE: 28-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 106010/95
 FILING DATE: 28-APR-1995
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: JP 106011/95
 FILING DATE: 28-APR-1995
 ATTORNEY/AGENT INFORMATION:
 NAME: Miller, Charles E.
 REGISTRATION NUMBER: 24,576
 REFERENCE/DOCKET NUMBER: 7426-043-999
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-8864/9741
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 9:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1048 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: genomic DNA
 ORGANISM: *Chlamydia pneumoniae*
 STRAIN: YK-41
 IMMEDIATE SOURCE:
 CLONE: 53-35
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 236 to 1012
 IDENTIFICATION METHOD: P
 US-08-809-326A-9
 Alignment Scores:
 Pred. No.: 1,95e-118 Length: 1048
 Score: 1262.00 Matches: 262
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 53,278 Indels: 0
 DB: 4 Gaps: 0
 US-09-889-314-2 (1-496) x US-08-809-326A-9 (1-1048)
 QY 1 AsphrsmksetserleserSerserSerglyProasphsnglnlysasnllwmetSer 20
 Db 227 GATCAACAACATGCTATTTCACTCTTTCAGACGCTGACAAATCAAAATATCATGCTC 286
 QY 21 GlnValleuthrserThrProglnglyValProglnglnAsplysleusSerglyAsnglu 40

Db 287 CAAGTTCGACATCGACACCCAGGCGCTGCCCAACAAGATAGCTGTCTGCAACGAA 346
 QY 41 ThrlysglnllleglnlnThrArglnlylysasnThrGlumetGluserAspAlaThr 60
 Db 347 ACGAAGCAAAATACAGCAAAACACGTCAGGGTAAAAACACTGAGATGAAAGGAGCCACT 406
 QY 61 lleaIaglyAlaSerGlylyAsplysThrserThrThrlysthrGluThrAlaPro 80
 Db 407 ATTGCTGGTGTCTTCTGAAAAGACAAACTTCTTGACTACAAAAACGAAACAGCTCCA 466
 QY 81 GlnGlnGlyValAlaAlaIaglylysgluserSergluserGlnlyAlaGlyAlaAspThr 100
 Db 467 CACACGAGAGTTCCTCTCGGAAGATCTCGAAGAGTCAAAAGCGGTGCTGATACT 526
 QY 101 GlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrlylleaMetGln 120
 Db 527 GGAGTATCAGACGCGCTGCTACTACAGATCAAAATCTGCAACAAAAATGCTATGCGAG 586
 QY 121 ThrSerllleglnlualaserlySerMetGluserThrLeuGluserLeuGlnSerleu 140
 Db 587 ACCTTAATGAAGGCGGACCAAAAGATGAGTCTTACCTTAGAGTCACTTCAAGCCCTC 646
 QY 141 SerAlaIaglnmetlysgluValGluAlaValAlaAlaIauserSerglylysser 160
 Db 647 AGTCGCCGCCAAATGAAGAAGATCGAAGCGGTGTGTGCTGCTCCACGAGAAAGT 706
 QY 161 SerGlySerAlaIySleuGlnThrProgluLeuProlySerglyValThrProArgSer 180
 Db 707 TCGGATCCGAAATTGGAAACCTGACCTCCCAAGCCGGGGTGACACCAAGATCA 766
 QY 181 GluValllleglnllleglyleuAlaAlaIySAlaIleGlnThrleuGlyGluAlaThr 200
 Db 767 GAGGTATCGAAATCGGACTCGCCCTGTCAACCAATTCAGACATTTGGAGAGCCACA 826
 QY 201 lysSerAlaIauserAsnThrlylaserThrGlnAlaGlnAlaAspGlnThrAsnlySleu 220
 Db 827 AAATCTGCTTATCTTAATCTATGCAAGTRACACACACAGACAGACCAAAATAAACA 886
 QY 221 GlyLeuGlnlysglnAlaIlelysIleaSplysgluArgGlnGlyTrglnGlnmetlyS 240
 Db 887 GGTCTAGAAAACCAAGCCATATAAATCGATAAAGCAGAAAGATCCAAAGAGATGAG 946
 QY 241 AlaAlaGlnGlnlySserlySpleuGlnGlyThrMetAspThrValasThrValMet 260
 Db 947 GCTGCCGACAGAGCTTAAGATCTCGAAGGAACAATGGATCTGCAATACTGTGATG 1006
 QY 261 lleaAla 262
 Db 1007 ATCGCG 1012
 RESULT 6
 US-08-809-326A-7
 ; Sequence 7, Application US/08809326A
 ; Patent No. 6165478
 ; GENERAL INFORMATION:
 ; APPLICANT: Izutsu, Hiroshi
 ; APPLICANT: Izutsu, Kazuhiko
 ; APPLICANT: Matsunoto, Akira
 ; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES, DNAS CODING
 ; TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS, TRANSFORMA
 ; TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMA
 ; TITLE OF INVENTION: CONTRAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION O
 ; TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREM
 ; TITLE OF INVENTION: OF ANTIBODY, METHOD AND AGENTS FOR DIAGNOSIS OF CHLAMYDIA
 ; TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND
 ; TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND
 ; TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
 ; TITLE OF INVENTION: PNEUMONIAE GENE
 ; NUMBER OF SEQUENCES: 31
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESS: Pennie & Edmonds LLP
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York

US-08-809-326A-25

Alignment Scores:

Pred. No.:	1,4e-117	Length:	5438
Score:	1265.00	Matches:	312
Percent Similarity:	62.39%	Conservative:	33
Best Local Similarity:	56.42%	Mismatches:	75
Query Match:	53.40%	Indels:	135
DB:	4	Gaps:	10

US-09-889-314-2 (1-496) x US-08-809-326A-25 (1-5438)

```

QY 4 MetSerLieserSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu 23
DB 540 ATGTCTATTTCATCTTCTTCAGAGCTGACATTCATAAAATATCATGTCTCAAGTTCG 599

QY 24 ThrSerThrProGlnGlyValProGlnAspLysLeuSerGlyValSerGln 43
DB 600 ACATGACACCCCGAGGCGCTGCCCAACAAGATAGCTCTGGCAACGAACGAAGCAA 659

QY 44 IleGlnIleThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGly 63
DB 660 ATACAGCAACACGTCAGGCTAATAAACACTGAGATGGAAGCGATGCCACTATGTGGT 719

QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGly 83
DB 720 GCTTCGGAAGAACAAACTCTCTGACTACAAAAACGAACAGCTCCACACAGCGGA 779

QY 84 ValAlaIleAlaGlyLysGlnSerSerGlnLysAlaGlyAlaAspThrGlyValSer 103
DB 780 GTTGGCTGGGGAAGAACTCTCAGAACTCAAAAGCGAGTGTGATCTGAGATATCA 839

QY 104 GlyAlaAlaIleAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle 123
DB 840 GAGCGGCTGCTACTACACACTCAATACTGCACAAAAATGTCTATGACACCTTAT 899

QY 124 GlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAla 143
DB 900 GAAGAGCGCAGCAAAAGTATGAGTCTACTTACCTTACCTTCAAAAGCCTCAGTCCGCG 959

QY 144 GlnMetLysGlnValGlnValValValAlaAlaIleLeuSerGlyLysSerSerGlySer 163
DB 960 CAAATCAAAAGAGCAACGCTGTGTGTGTGCTGCTCAGGGAAGATTCGGGTTCC 1019

QY 164 AlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnValIle 183
DB 1020 GCAAAATTTGAAACACCTAGCTCCCAAGCCCGGGTACACACAGATCAGAGTTATC 1079

QY 184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlnValAlaThrLysSerAla 203
DB 1080 GAATCGGACTCGCGCTGCTTAAGCAATTCAGACATTTGGAGAAAGCCCAAAATCTGCC 1139

QY 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnLeuGlu 223
DB 1140 TTATCTAATCTAGCAAGTACACAGCAAGCAGCAACCAAAATTAAGTCTAGTACGAA 1199

QY 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGluThrGlnGlnMetLysAlaAlaGlu 243
DB 1200 AAGCAAGCATTAAAAATCATTAAGACAGAGAAAGATACCAAGATGAGAGGCTCCGGA 1259

QY 244 GlnLysSerLysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaVal 263
DB 1260 CAGAAGTCTTAAGATCTCAGAGAACATGATGATCTCAATACCTGTGATGATCGGAA- 1318

QY 264 SerValAlaIleThrVal-----IleSerIleValAla 274
DB 1319 GGGGTTTCAATTGGCCATGGGGCCCTTAATTAATTAATCTGAGAGATCCAGATCTATG 1378

QY 275 AlaIlePheThrCysGlyLeuAlaGlyLeu-----AlaAlaGlyAlaVal 292
DB 1379 ATGATCTCTACGCGCAGCATCTGTGGCGGATCACCGCGGCCACAGTGGCTGCT 1438

QY 293 GlyAlaAla----- 295

```

```

DB 1439 GCGCGCTA-TATCGCGCATCACCAGATGGGGAAGATCGGGCTGCCACTTCGGGCTCAT 1497
QY 296 -----AlaIleGlyGlyAlaIleAlaVal 303
DB 1498 GAGCGCTGTTTCGGCGGTGATGTGGCAGCGCCGTGGCGGGGACTGTTGGGCGCC 1557
QY 304 -----AlaAlaAla----- 306
DB 1558 ATCTCTTGACATGACACCATCTCTTGGCGGGGCTGTCAACGGCTCAACTACTACTG 1617
QY 306 ----- 306
DB 1618 GCGCTGCTCTAATGACAGAGTCGATTAAGAGAGCGTCGACGATGCCCTTGAGAGCC 1677
QY 307 -----ThrValAlaThrGlnIle 313
DB 1678 TTCAACCCAGTCAGCTCTTCGGTGGCGCGGACATGATGCTGCGCCACTATG 1737
QY 314 ThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValArgGlnAla 333
DB 1738 ACTGCTCTTATCATGACACTCGTAGACAG----- 1770
QY 334 IleThrAlaAlaIle-----LysAlaAlaValLys 343
DB 1771 GTCGCGGAGCGGCTCGGTGATTTTCGGCAGAGACCGCTTCGCTGGAGCGCGCATG 1830
QY 344 SerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerLys 363
DB 1831 ATCGGCTGCTCCCTTCGGTATTCGGAATCTTCGACGCGCTTCGCAACCTTCCTCACT 1890
QY 364 GlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAspPheProLysLeu 383
DB 1891 GGTCCCGCC-----ACCAACGTTTCGCGCAGAAAG 1920
QY 384 SerLysValIleSerSerLeu----- 390
DB 1921 CAGCGCATATCGCGGATGCGCGCCGACGCGCTGAGCTGCTGCTGCGCTTCGCG 1980
QY 391 ThrSerLysThrValThrValGlyValAlaValAlaIleAlaProAlaLeuGlnLys 410
DB 1981 ACGGAGGCTGGATGCTTCCTCCCATTTATGATTTCTGCTTCGGCGCGCATCGGATG 2040
QY 411 GlyIleMetGlnLeuSerGlnMetGlnAsnValAlaGlnPheGlnLysGlu 430
DB 2041 CCGCGGTTCCAGGCCATGTGTTCACAGCGAGTAGATGACACCATCAGGACAGCTTCAA 2100
QY 431 ValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMet 443
DB 2101 ---GGATCGCTCGCGGCTCTTACCAAGCTTAACCTGATC 2136

RESULT 5
US-08-809-326A-9
; Sequence 9, Application US/08809326A
; Patent No. 6165478
; GENERAL INFORMATION:
; APPLICANT: Iizutsu, Hiroshi
; APPLICANT: Obara, Kazuhiko
; APPLICANT: Matsumoto, Akira
; TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES, DNAS CODING
; TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING
; TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMA
; TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION O
; TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREM
; TITLE OF INVENTION: OF ANTIBODY, METHOD AND REAGENTS FOR DETECTION OF CHLAMYDIA
; TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND
; TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Penile & Edmonds LLP
; STREET: 1155 Avenue of the Americas

```

Db 1231 TCGGTTCCGCAAAATTTGGAAACACCTGAGCTCCCAAGCCCGGGGTGACACCAAGATCA 1290
Qy 181 GlnValIleGluIleGlyLeuAlaLeuAlaValAlaIleGlnThrLeuGlyAlaThr 200
Db 1291 GAGGTTATGCAAAATTCGACCTCGGCTTGAAGCAATTCAGACATTCGGAGAACCCACA 1350
Qy 201 LysSerAlaLeuSerAsnTyrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeu 220
Db 1351 AAATCTCCCTTATCTAATCACTATGCAAGTACCAAGCACAAGCACAACCAATTAACCTA 1410
Qy 221 GlyLeuLysGlnAlaIleLysIleAspLysGlnArgGlnGluLysGlnGluLys 240
Db 1411 GGTCTAAGAAAGCAAGCGATTAATGATGAAGCAAGCAAGCAAGCAAGCAAGATGAAG 1470
Qy 241 AlaAlaGlnLysSerLysAspLeuGlyLysThrMetLysPheValAsnThrValMet 260
Db 1471 GCTCCGACAGAGAGTCTAAAGATCTGAAGCAATGATCTGCTCATCTACTGTGATG 1530
Qy 261 IleAlaValSerValAlaIleThrVal-----LysSer 271
Db 1531 ATCCGGAA-GGGGTTCGATTCGATGGGGCCCTTAATTAATTAATCAAGAGATCCAG 1589
Qy 272 IleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeu-----AlaAlaGly 289
Db 1590 ATCTAATGATGATCTCTACGCGGACGATCTGCGCGGACATCCAGCGCCACAGGT 1649
Qy 290 AlaAlaValGlyAlaAla----- 295
Db 1650 GCGGTTGCTGGCGCCTA-TATCGCCGACATACCGATGGGGAAGTGGGGCTGCCACTT 1708
Qy 296 -----AlaAlaGlyGlyAla 300
Db 1709 CCGGCTCATGACGCTTGTTCGGGTGGATGGTGGCAGGCGCGGCGGAGCTG 1768
Qy 301 AlaGlyAla-----AlaAlaAla----- 306
Db 1769 TTGGGGCCATCTCTTGCATGACCACTTCTTGGCGGGCGGTCTCAAGGGCTCAAC 1828
Qy 306 ----- 306
Db 1829 CTACTAGTGGGCTCTCTTAATGACAGATGCGCATTAAGGAGAGCGTCGACCGATGCC 1888
Qy 307 -----ThrValAla 310
Db 1889 TTGAGAGCCTTCAACCCAGTACCTCTTCGGTGGCGCGGCGGATGACTCTGTCGCC 1948
Qy 311 ThrGlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaVal 330
Db 1949 GCACCTATGACTGTCTTCTTATCATGACAACTGTAGACAG----- 1990
Qy 331 ArgGlnAlaIleThrAlaAlaIle-----LysAla 340
Db 1991 -----TGCGCGGACGCGCTCTGGTCAATTTGGCGGAGACCGCTTTCCTGGAGC 2041
Qy 341 AlaValLysSerGlyLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAla 360
Db 2042 GCGACGATGATGCGCTGCTGCGGTATGCGATCTTGACGCGCTCGCTCAAGCC 2101
Qy 361 IleSerLysGlyLysSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPhe 380
Db 2102 TTTCGCACTGCTCCGCC-----ACCAACGCTTC 2131
Qy 381 ProLysLeuSerLysValIleSerSerLeu----- 390
Db 2132 GCGGAGAAAGCAAGCCATTATGCGCGCATGGCGGCGCGGCTAGCTTCTGCTG 2191
Qy 391 -----ThrSerLysTrpValThrValGlyValGlyValAlaValAlaAlaProAla 407
Db 2192 GCGTTCCGAGCGAGCGGCTGATGCGCTTCCCATTAAGATCTTCGCTCGCGCGC 2251
Qy 408 LeuGlyLysGlyIleMetGlnMetLysSerGlnMetGlnGlnAsnValAlaGlnPhe 427
Db 2252 ATCGGATGCCCGCGTTGACGCGCATGCTGTCCAGGCGAGTATGACGACACATCAGGGA 2311

Qy 428 GlnLysGlnValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMet 443
Db 2312 CAGCTTCAA---GGATCGCTCGCGCTCTTACCAAGCTTAATTCGATC 2356
RESULT 4
US-08-809-326A-25
Sequence 25, Application US/08809326A
Patent No. 6165478
GENERAL INFORMATION:
APPLICANT: Izutsu, Hiroshi
APPLICANT: Matsumoto, Akira
APPLICANT: Obara, Kazuhiko
TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES, DNAS CODING
THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMA
TIONS OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS CODING
THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMA
TIONS OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT
OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND
MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND
REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
PNEUMONIAE GENE
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESSES:
ADDRESSER: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,326A
FILING DATE: 19-MAR-1997
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 224711/94
FILING DATE: 20-SEP-1994
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 106006/95
FILING DATE: 28-APR-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 106008/95
FILING DATE: 28-APR-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 106009/95
FILING DATE: 28-APR-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 106010/95
FILING DATE: 28-APR-1995
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: JP 106011/95
FILING DATE: 28-APR-1995
ATTORNEY/AGENT INFORMATION:
NAME: Miller, Charles E.
REGISTRATION NUMBER: 24,576
REFERENCE/DOCKET NUMBER: 7426-043-999
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864/9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 25:
SEQUENCE CHARACTERISTICS:
LENGTH: 5438 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: Other nucleic acid; Plasmid


```

1 ? Sequence 17 Application US/08809326A
2 ? Patent No. 6165478
3 ? GENERAL INFORMATION:
4 ? APPLICANT: Iizutsu, Hiroshi
5 ? APPLICANT: Obata, Kazuhiko
6 ? APPLICANT: Matsumoto, Akira
7 ? TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES, DNAS CODING
8 ? TITLE OF INVENTION: FUSED PROTEINS CONTAINING THE POLYPEPTIDES, DNAS, TRANSFORMANTS
9 ? TITLE OF INVENTION: THEREFOR, RECOMBINANT VECTORS CARRYING THE POLYPEPTIDES, DNAS, TRANSFORMANTS
10 ? TITLE OF INVENTION: CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF
11 ? TITLE OF INVENTION: ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT
12 ? TITLE OF INVENTION: PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND/OR
13 ? TITLE OF INVENTION: MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND
14 ? TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA
15 ? NUMBER OF SEQUENCES: 31
16 ? PNEUMONIAE GENE
17 ? CORRESPONDENCE ADDRESS:
18 ? ADDRESSEE: Pennile & Edmonds LLP
19 ? STREET: 1155 Avenue of the Americas
20 ? CITY: New York
21 ? STATE: New York
22 ? COUNTRY: USA
23 ? ZIP: 10036-2711
24 ? COMPUTER READABLE FORM:
25 ? MEDIUM TYPE: Floppy disk
26 ? COMPUTER: IBM PC compatible
27 ? OPERATING SYSTEM: PC-DOS/MS-DOS
28 ? SOFTWARE: PatentIn Release #1.0, Version #1.30
29 ? CURRENT APPLICATION DATA:
30 ? APPLICATION NUMBER: US/08/809,326A
31 ? FILING DATE: 19-MAR-1997
32 ? CLASSIFICATION: 435
33 ? PRIOR APPLICATION DATA:
34 ? APPLICATION NUMBER: JP 224711/94
35 ? FILING DATE: 20-SEP-1994
36 ? PRIOR APPLICATION DATA:
37 ? APPLICATION NUMBER: JP 106006/95
38 ? FILING DATE: 28-APR-1995
39 ? PRIOR APPLICATION DATA:
40 ? APPLICATION NUMBER: JP 106008/95
41 ? FILING DATE: 28-APR-1995
42 ? PRIOR APPLICATION DATA:
43 ? APPLICATION NUMBER: JP 106009/95
44 ? FILING DATE: 28-APR-1995
45 ? PRIOR APPLICATION DATA:
46 ? APPLICATION NUMBER: JP 106010/95
47 ? FILING DATE: 28-APR-1995
48 ? PRIOR APPLICATION DATA:
49 ? APPLICATION NUMBER: JP 106011/95
50 ? FILING DATE: 28-APR-1995
51 ? ATTORNEY/AGENT INFORMATION:
52 ? NAME: Miller, Charles E.
53 ? REGISTRATION NUMBER: 24,576
54 ? REFERENCE/DOCKET NUMBER: 7426-043-999
55 ? TELECOMMUNICATION INFORMATION:
56 ? TELEPHONE: (212) 790-9090
57 ? TELEFAX: (212) 869-8864/9741
58 ? TELEX: 66141 PENNILE
59 ? INFORMATION FOR SEQ. ID NO: 17:
60 ? SEQUENCE CHARACTERISTICS:
61 ? LENGTH: 1947 base pairs
62 ? TYPE: nucleic acid
63 ? STRANDEDNESS: double
64 ? TOPOLOGY: linear
65 ? MOLECULE TYPE: other nucleic acid; synthetic DNA
66 ?
67 ?
68 ?
69 ?
70 ?
71 ?
72 ?
73 ?
74 ?
75 ?
76 ?
77 ?
78 ?
79 ?
80 ?
81 ?
82 ?
83 ?
84 ?
85 ?
86 ?
87 ?
88 ?
89 ?
90 ?
91 ?
92 ?
93 ?
94 ?
95 ?
96 ?
97 ?
98 ?
99 ?
100 ?
101 ?
102 ?
103 ?
104 ?
105 ?
106 ?
107 ?
108 ?
109 ?
110 ?
111 ?
112 ?
113 ?
114 ?
115 ?
116 ?
117 ?
118 ?
119 ?
120 ?
121 ?
122 ?
123 ?
124 ?
125 ?
126 ?
127 ?
128 ?
129 ?
130 ?
131 ?
132 ?
133 ?
134 ?
135 ?
136 ?
137 ?
138 ?
139 ?
140 ?
141 ?
142 ?
143 ?
144 ?
145 ?
146 ?
147 ?
148 ?
149 ?
150 ?
151 ?
152 ?
153 ?
154 ?
155 ?
156 ?
157 ?
158 ?
159 ?
160 ?
161 ?
162 ?
163 ?
164 ?
165 ?
166 ?
167 ?
168 ?
169 ?
170 ?
171 ?
172 ?
173 ?
174 ?
175 ?
176 ?
177 ?
178 ?
179 ?
180 ?
181 ?
182 ?
183 ?
184 ?
185 ?
186 ?
187 ?
188 ?
189 ?
190 ?
191 ?
192 ?
193 ?
194 ?
195 ?
196 ?
197 ?
198 ?
199 ?
200 ?
201 ?
202 ?
203 ?
204 ?
205 ?
206 ?
207 ?
208 ?
209 ?
210 ?
211 ?
212 ?
213 ?
214 ?
215 ?
216 ?
217 ?
218 ?
219 ?
220 ?
221 ?
222 ?
223 ?
224 ?
225 ?
226 ?
227 ?
228 ?
229 ?
230 ?
231 ?
232 ?
233 ?
234 ?
235 ?
236 ?
237 ?
238 ?
239 ?
240 ?
241 ?
242 ?
243 ?
244 ?
245 ?
246 ?
247 ?
248 ?
249 ?
250 ?
251 ?
252 ?
253 ?
254 ?
255 ?
256 ?
257 ?
258 ?
259 ?
260 ?
261 ?
262 ?
263 ?
264 ?
265 ?
266 ?
267 ?
268 ?
269 ?
270 ?
271 ?
272 ?
273 ?
274 ?
275 ?
276 ?
277 ?
278 ?
279 ?
280 ?
281 ?
282 ?
283 ?
284 ?
285 ?
286 ?
287 ?
288 ?
289 ?
290 ?
291 ?
292 ?
293 ?
294 ?
295 ?
296 ?
297 ?
298 ?
299 ?
300 ?
301 ?
302 ?
303 ?
304 ?
305 ?
306 ?
307 ?
308 ?
309 ?
310 ?
311 ?
312 ?
313 ?
314 ?
315 ?
316 ?
317 ?
318 ?
319 ?
320 ?
321 ?
322 ?
323 ?
324 ?
325 ?
326 ?
327 ?
328 ?
329 ?
330 ?
331 ?
332 ?
333 ?
334 ?
335 ?
336 ?
337 ?
338 ?
339 ?
340 ?
341 ?
342 ?
343 ?
344 ?
345 ?
346 ?
347 ?
348 ?
349 ?
350 ?
351 ?
352 ?
353 ?
354 ?
355 ?
356 ?
357 ?
358 ?
359 ?
360 ?
361 ?
362 ?
363 ?
364 ?
365 ?
366 ?
367 ?
368 ?
369 ?
370 ?
371 ?
372 ?
373 ?
374 ?
375 ?
376 ?
377 ?
378 ?
379 ?
380 ?
381 ?
382 ?
383 ?
384 ?
385 ?
386 ?
387 ?
388 ?
389 ?
390 ?
391 ?
392 ?
393 ?
394 ?
395 ?
396 ?
397 ?
398 ?
399 ?
400 ?
401 ?
402 ?
403 ?
404 ?
405 ?
406 ?
407 ?
408 ?
409 ?
410 ?
411 ?
412 ?
413 ?
414 ?
415 ?
416 ?
417 ?
418 ?
419 ?
420 ?
421 ?
422 ?
423 ?
424 ?
425 ?
426 ?
427 ?
428 ?
429 ?
430 ?
431 ?
432 ?
433 ?
434 ?
435 ?
436 ?
437 ?
438 ?
439 ?
440 ?
441 ?
442 ?
443 ?
444 ?
445 ?
446 ?
447 ?
448 ?
449 ?
450 ?
451 ?
452 ?
453 ?
454 ?
455 ?
456 ?
457 ?
458 ?
459 ?
460 ?
461 ?
462 ?
463 ?
464 ?
465 ?
466 ?
467 ?
468 ?
469 ?
470 ?
471 ?
472 ?
473 ?
474 ?
475 ?
476 ?
477 ?
478 ?
479 ?
480 ?
481 ?
482 ?
483 ?
484 ?
485 ?
486 ?
487 ?
488 ?
489 ?
490 ?
491 ?
492 ?
493 ?
494 ?
495 ?
496 ?
497 ?
498 ?
499 ?
500 ?
501 ?
502 ?
503 ?
504 ?
505 ?
506 ?
507 ?
508 ?
509 ?
510 ?
511 ?
512 ?
513 ?
514 ?
515 ?
516 ?
517 ?
518 ?
519 ?
520 ?
521 ?
522 ?
523 ?
524 ?
525 ?
526 ?
527 ?
528 ?
529 ?
530 ?
531 ?
532 ?
533 ?
534 ?
535 ?
536 ?
537 ?
538 ?
539 ?
540 ?
541 ?
542 ?
543 ?
544 ?
545 ?
546 ?
547 ?
548 ?
549 ?
550 ?
551 ?
552 ?
553 ?
554 ?
555 ?
556 ?
557 ?
558 ?
559 ?
560 ?
561 ?
562 ?
563 ?
564 ?
565 ?
566 ?
567 ?
568 ?
569 ?
570 ?
571 ?
572 ?
573 ?
574 ?
575 ?
576 ?
577 ?
578 ?
579 ?
580 ?
581 ?
582 ?
583 ?
584 ?
585 ?
586 ?
587 ?
588 ?
589 ?
590 ?
591 ?
592 ?
593 ?
594 ?
595 ?
596 ?
597 ?
598 ?
599 ?
600 ?
601 ?
602 ?
603 ?
604 ?
605 ?
606 ?
607 ?
608 ?
609 ?
610 ?
611 ?
612 ?
613 ?
614 ?
615 ?
616 ?
617 ?
618 ?
619 ?
620 ?
621 ?
622 ?
623 ?
624 ?
625 ?
626 ?
627 ?
628 ?
629 ?
630 ?
631 ?
632 ?
633 ?
634 ?
635 ?
636 ?
637 ?
638 ?
639 ?
640 ?
641 ?
642 ?
643 ?
644 ?
645 ?
646 ?
647 ?
648 ?
649 ?
650 ?
651 ?
652 ?
653 ?
654 ?
655 ?
656 ?
657 ?
658 ?
659 ?
660 ?
661 ?
662 ?
663 ?
664 ?
665 ?
666 ?
667 ?
668 ?
669 ?
670 ?
671 ?
672 ?
673 ?
674 ?
675 ?
676 ?
677 ?
678 ?
679 ?
680 ?
681 ?
682 ?
683 ?
68
```

Query Match:	97.76%	Indels:	0
DB:	4	Gaps:	0
US-09-889-314-2 (1-496) x US-08-809-326A-17 (1-1947)			
OY	4 MetSerIleSerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeu	23	
DB	484 ATGTCTATTTCATCTCTTCAGACCTGACAAATCAAAAAATATCATATGCTCAAGCTTCG	543	
OY	24 ThrSerThrProGlnGlyValProGlnGlnAspLysSerGlyLysnsluThrLysGln	43	
DB	544 ACATGCACACCCCAAGGGGTGCCCAACAGATAGCTGCTGTGGCAACGAACGAACAA	603	
OY	44 IleGlnGlnThrArgGlnGlnGlyLysAsnThrGluMetGlnSerAspAlaThrIleAagly	63	
DB	604 ATACACCAACACGTCAGGGTAAAAACACTGATGAGTAAGAACGATGCGCATATGCTGGGT	663	
OY	64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGly	83	
DB	664 GCTTCGAAAAAGACAAACCTTCGACTACAAAACAGAAACAGCTCCACAACAGGGA	723	
OY	84 ValAlaIleGlyLysGlnSerSerGlnSerGlnLysAlaGlyAlaAspThrGlyValSer	103	
DB	724 GTTGCTCTCTGGGAAAGAAATCCTCAGAAAGTCAAAAGCAGGCTGCTGATCTGAGATCA	783	
OY	104 GlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIle	123	
DB	784 GGACCGGCTCTCTACTACAGCATCAAAATCTGCACAAAATTCGTTAGCAAGCTTATT	843	
OY	124 GlnGlnLysSerLysSerMetGlnSerThrIleGlnGlnSerGlnSerLysSerAlaAla	143	
DB	844 GAAGGGCGGACCAAAATATGAGAGTCTACCTTAGAGTCACTTCAAAAGCCTCAGTCCGGC	903	
OY	144 GlnMetLysGlnValGlnAlaValValValAlaAlaIleSerGlyLysSerSerGlySer	163	
DB	904 CAATGGAAGGAGTCGAAAGGGGTGTGTGTGCTGCTCCCTCTCAGGGAAAAAGTTGGGTTC	963	
OY	164 MetLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnValIle	183	
DB	964 GCAAAATTGGAAACACCTGAGACTCCCAACCCGGGGGTGACCAAGATCAAGGTATTC	1023	
OY	184 GlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlnLysAlaThrLysSerAla	203	
DB	1024 GAATTCGCGCTCGCCCTGTGAAGCAATTCAGCATTTGGGAGAACCCCAAAATCTGCC	1083	
OY	204 LeuSerAsnThrLysSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnGlnGln	223	
DB	1084 TTATCTAATATGCAAGTACACAGACCAACCAACCAAAATAACTAAGGTCTAGCA	1143	
OY	224 LysGlnAlaIleLysLileAspLysGlnLysGlnGlnGlnGlnGlnGlnGlnGlnGlnGln	243	
DB	1144 AACCAAGCATAAAAATGATTAACAGACGACGAAAGATTAACAAAGGATGAAGGCTGCCGA	1203	
OY	244 GlnLysSerLysAspLeuGlnGlnThrMetAspThrValAsnThrValMetIleAlaVal	263	
DB	1204 CAGACGCTAAATCTCGAAGGAAACAATGATCTGTCAATCTGATGATGATCGCGTT	1263	
OY	264 SerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeu	283	
DB	1264 TCTGTTCCCTTACAGTATTTCTATTGTGCTGCTATTATTACATGGGAGCTGCACTC	1323	
OY	284 AlaGlyLeuAlaIleGlyAlaAlaValGlyAlaAlaAlaAlaIleGlyAlaAlaIleGlyAla	303	
DB	1324 GCTGACTCGCTGGGGAGCTGCTGATGAGTGCAGCGGCAAGCTGAGGTGACGAGAGCT	1383	
OY	304 AlaAlaAlaThrThrValAlaThrGlnIleThrValGlnAlaAlaValGlnAlaValLys	323	
DB	1384 GCTGCGGCAACACGAGTGAACACAAATTCAGTTCAAGCTGTGTCCAAAGCGGTGAA	1443	
OY	324 GlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLys	343	
DB	1444 CAACTGTTATTCACAGGTGTCAGCAAGCATACCGCGGCTATTAAGCGGCTCTCAA	1503	

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: January 27, 2003, 14:11:52 ; Search time 51 Seconds
(without alignments)
2982.582 Million cells updates/sec

Title: US-09-889-314-2

perfect score: 2369
Sequence: 1 DTNMSISSSGPDNQRKIMS.....LKAVALISGALGAKTKTNF 496

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 441362 seqs, 15338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ .p2n.model -DEV=rlh
-O=/cgn2_1/USPTO.spool/US09889314/runat.24012003.144541.10130/app.query.fasta.1.647
-DB=Issued Patents.NA -QFMT=fastap -SUFFIX=trn1 -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdl
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=45
-MODE=LOCAL -OUTFMT=ptc -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09889314.ecgn_1.1.17_gunat.24012003.144541.10130 -NCPU=6 -ICPU=3
-NO_XIPYX -NO_MMAB -LARGEOUTERT -NEG_SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=7
-WARN.TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued Patents.NA.*
1: /cgn2_6/ptodata/1/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2316	97.8	1464	4 US-08-809-326A-3	Sequence 3, Appl 1
2	2316	97.8	1947	4 US-08-809-326A-17	Sequence 17, Appl 1
3	1282	54.1	5658	4 US-08-809-326A-10	Sequence 10, Appl 1
4	1265	53.4	5438	4 US-08-809-326A-25	Sequence 25, Appl 1
5	1262	53.3	1048	4 US-08-809-326A-9	Sequence 9, Appl 1
6	1245	52.6	777	4 US-08-809-326A-7	Sequence 7, Appl 1
7	1245	52.6	813	4 US-08-809-326A-4	Sequence 4, Appl 1
8	1245	52.6	1296	4 US-08-809-326A-18	Sequence 18, Appl 1
9	220.5	9.3	5393	2 US-08-591-079-9	Sequence 9, Appl 1
10	218.5	9.2	3622	2 US-08-591-079-7	Sequence 7, Appl 1
11	213.5	9.0	32768	4 US-08-961-527-71	Sequence 71, Appl 1
12	203	8.6	46819	4 US-09-453-702B-72	Sequence 72, Appl 1

13	199	8.4	4483	4 US-08-961-527-363	Sequence 363, App
14	193	8.1	1702	4 US-08-961-527-367	Sequence 367, App
15	192	8.1	2550	4 US-08-961-527-364	Sequence 364, App
16	188	7.9	1436	4 US-08-961-527-365	Sequence 365, App
17	186	7.9	7766	4 US-09-125-619-3	Sequence 3, Appl 1
18	185	7.8	1743	2 US-08-591-079-1	Sequence 1, Appl 1
19	172.5	7.3	869	4 US-08-961-527-369	Sequence 369, App
20	170	7.2	43360	4 US-09-453-702B-206	Sequence 206, App
21	170	7.2	45355	4 US-09-453-702B-261	Sequence 261, App
22	167	7.0	3744	4 US-08-961-527-368	Sequence 368, App
23	166.5	7.0	941	4 US-08-961-527-370	Sequence 370, App
24	165.5	7.0	750	4 US-08-961-527-372	Sequence 372, App
25	158.5	6.7	807	4 US-08-961-527-373	Sequence 373, App
26	157	6.6	1068	4 US-08-961-527-374	Sequence 374, App
27	155.5	6.6	4411529	4 US-09-103-840A-1	Sequence 1, Appl 1
28	155	6.5	4403765	4 US-09-103-840A-2	Sequence 2, Appl 1
29	152.5	6.4	3106	2 US-08-591-079-5	Sequence 5, Appl 1
30	152	6.4	6530	2 US-08-146-930-1	Sequence 1, Appl 1
31	152	6.4	6530	3 US-08-458-240-1	Sequence 1, Appl 1
32	152	6.4	6530	5 PCT-US93-03993-1	Sequence 1, Appl 1
33	151	6.4	2835	4 US-09-134-001C-1515	Sequence 1515, App
34	151	6.4	6414	4 US-09-134-001C-1626	Sequence 1626, App
35	150.5	6.4	3561	4 US-09-134-001C-1685	Sequence 1685, App
36	150	6.3	5118	4 US-08-669-785-3	Sequence 3, Appl 1
37	148	6.2	6530	2 US-08-146-930-1	Sequence 1, Appl 1
38	148	6.2	6530	3 US-08-458-240-1	Sequence 1, Appl 1
39	148	6.2	6530	5 PCT-US93-03993-1	Sequence 1, Appl 1
40	147.5	6.2	690	4 US-08-961-527-377	Sequence 377, App
41	147.5	6.2	4649	6 US-08-669-785-1	Sequence 1, Appl 1
42	147.5	6.2	6441	4 US-08-669-785-1	Sequence 1, Appl 1
43	147.5	6.2	6441	6 US-08-669-785-1	Sequence 1, Appl 1
44	147.5	6.2	8078	4 US-09-702-251-3	Sequence 3, Appl 1
45	146.5	6.2	3543	4 US-09-206-942-64	Sequence 64, Appl 1

ALIGNMENTS

RESULT 1
US-08-809-326A-3
Sequence 3, Application US/08809326A
Patent No. 6165478
GENERAL INFORMATION:
APPLICANT: Iizutsu, Hiroshi
APPLICANT: Obata, Kazuhiko
APPLICANT: Matsumoto, Akira
TITLE OF INVENTION: CHLAMYDIA PNEUMONIAE ANTIGENIC POLYPEPTIDES, DNAS CODING THEREFOR, RECOMBINANT VECTORS CARRYING THE DNAS, TRANSFORMANTS CONTAINING THE RECOMBINANT VECTORS, METHOD FOR PRODUCTION OF ANTIBODY, METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT OF ANTIBODY, METHOD AND REAGENTS FOR DIAGNOSIS OF CHLAMYDIA PNEUMONIAE INFECTIONS, PROBES AND PRIMERS FOR DETECTION AND MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE, AND METHOD AND REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE
TITLE OF INVENTION: REAGENTS FOR DETECTION AND/OR MEASUREMENT OF CHLAMYDIA PNEUMONIAE GENE
TITLE OF INVENTION: PNEUMONIAE GENE
NUMBER OF SEQUENCES: 31
CORRESPONDENCE ADDRESSES:
ADDRESSEE: Penile & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/809,326A
FILING DATE: 19-MAR-1997
CLASSIFICATION: 435
PRIOR APPLICATION DATA:

QY 294 AAAAGCAA 301
|||||||
DB 25 AAAAGCAA 32

RESULT 42

E88637
protein W09G12.6 [imported] - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 10-May-2001
C:Accession: E88637
R:Anonymous, The C. elegans Sequencing Consortium.
S:Genbank, 2012-2018, 1998
A:Title: Genome sequence of the nematode C. elegans: a platform for investigating biolog
A:Reference number: A75000; MID:99069613; PMID:9851916
A:Note: see websites genome.wustl.edu/gsc/C_elegans/ and www.sanger.ac.uk/Projects/C_ele
A:Note: published errata appeared in Science 283, 35, 1999; Science 283, 2103, 1999; and
A:Accession: E88637
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-178 <STO>
A:Cross-references: GB:chr_IV; PIDN:AAC04448.1; PID:92911883; GSPDB:GN00022; CESP:W09G12
C:Genetics:
A:Gene: W09G12.6
A:Map position: 4

Query Match 1.6%; Score 8; DB 2; Length 178;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 AAGGAGAA 303
|||||||
DB 124 AAGGAGAA 131

RESULT 43

A85217
hypothetical protein AT4g19200 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
C:Accession: A85217
R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Sprin
Nature 402, 769-777, 1999
A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
A:Reference number: A85001; MID:20083448; PMID:10617198
A:Accession: A85217
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-179 <STO>
A:Cross-references: GB:NC_001268; MID:q7268715; PIDN:CAB78922.1; GSPDB:GN00140
C:Genetics:
A:Gene: AT4g19200
A:Map position: 4

Query Match 1.6%; Score 8; DB 2; Length 179;
Best Local Similarity 100.0%; Pred. No. 20;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 299 GAAGAAAA 306
|||||||
DB 99 GAAGAAAA 106

RESULT 44

E70536
hypothetical protein RV0804 - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 22-Oct-1999
C:Accession: E70536
R:Cole, S.T.; Brosch, R.; Parhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.
; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998

A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno
A:Reference number: A70500; MID:9825987; PMID:9634230
A:Accession: E70536
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-209 <COL>
A:Cross-references: GB:Z95618; GB:AL123456; MID:q3261788; PIDN:CAB09107.1; PID:631727
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: RV0804

Query Match 1.6%; Score 8; DB 2; Length 209;
Best Local Similarity 100.0%; Pred. No. 23;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 300 AAGAAAT 307
|||||||
DB 63 AAGAAAT 70

RESULT 45

G97631
hypothetical protein AGR_C_4120 [imported] - Agrobacterium tumefaciens (strain C58, C
C:Species: Agrobacterium tumefaciens
C:Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 11-Jan-2002
C:Accession: G97631
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm
A.; Liu, F.; Wollam, C.; Allinger, M.; Doughly, D.; Scott, C.; Lappas, C.; Markelz,
Science 294, 2323-2328, 2001
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium
A:Reference number: A97359; PMID:11743194
A:Accession: G97631
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-214 <KUR>
A:Cross-references: GB:AE007869; PIDN:AAK88008.1; PID:915157423; GSPDB:GN00169
C:Genetics:
A:Gene: AGR_C_4120
A:Map position: circular chromosome

Query Match 1.6%; Score 8; DB 2; Length 214;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 284 AGAAGAA 291
|||||||
DB 165 AGAAGAA 172

Search completed: January 27, 2003, 16:37:07
Job time : 53 secs

R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.; M.; Shen, M.; Yamahewyan, J.J.; Lam, P.; McDonald, L.; Uterback, T.; Zaleski, C.; Ma, S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.
 Science 286, 1571-1577, 1999
 A:Title: Genome sequence of the radioresistant bacterium *Deinococcus radiodurans* R1.
 A:Reference number: A75250; MUID:20036896; PMID:10567266
 A:Accession: H75262
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-145 <WMT>
 A:Cross-references: GB:AE002082; GB:AE000513; NID:96460347; PIDN:AAF12073.1; PID:9646035
 A:Experimental source: strain R1
 C:Genetics:
 A:Gene: DR2529
 A:Map position: 1

Query Match 1.6%; Score 8; DB 2; Length 145;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 284 GLAGAAG 291
 DB 33 GAGAGAA 40
 |||||

RESULT 38
 S43893
 H+-exporting ATPase (EC 3.6.3.6) lipid-binding protein - *Neurospora crassa*
 C:Species: *Neurospora crassa*
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
 C:Accession: S43893
 R:Sista, H.; Wechsler, M.A.; Bowman, B.J.
 Mol. Gen. Genet. 243, 82-90, 1994
 A:Title: The proteolipid subunit of the *Neurospora crassa* vacuolar ATPase: isolation of
 A:Reference number: S43893; MUID:94247360; PMID:8190074
 A:Accession: S43893
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-161 <STS>
 A:Cross-references: EMBL:L07105; NID:q168929; PIDN:AAA19974.1; PID:q168930
 A:Note: the authors translated the codon ATC for residue 63 as Tyr and GCT for residue 1
 C:Genetics:
 A:Introns: 5/2; 9/3; 53/1; 141/1
 C:Superfamily: vacuolar H+-transporting ATPase 16k chain
 C:Keywords: ATP; hydrolase

Query Match 1.6%; Score 8; DB 1; Length 161;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 282 GAGAGAG 289
 DB 99 GAGAGAG 106
 |||||

RESULT 39
 S32970
 H+-exporting ATPase (EC 3.6.3.6) chain c-6, vacuolar - fission yeast (*Schizosaccharomyces*
 N:Alternate names: vacuolar H+-transporting ATPase c-6
 C:Species: *Schizosaccharomyces pombe*
 C>Date: 19-Mar-1997 #sequence_revision 25-Apr-1997 #text_change 03-Jun-2002
 C:Accession: S32970; J38033
 R:Toyama, R.; Goldstein, D.J.; Schlegel, R.; Dhar, R.
 Yeast 7, 989-991, 1991
 A:Title: A genomic sequence of the *Schizosaccharomyces pombe* 16 kDa vacuolar H(+)-ATPase
 A:Reference number: S32970; MUID:92206078; PMID:1839480
 A:Accession: S32970
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-161 <TOY>
 A:Cross-references: EMBL:X59947; NID:q5135; PIDN:CAA42572.1; PID:q5136
 R:Murphy, L.; Harris, D.; Barrell, B.G.; Rajandream, M.A.; Wood, V.
 submitted to the EMBL Data Library, August 1997
 A:Reference number: Z21763

A:Accession: T38033
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-161 <MUR>
 A:Cross-references: EMBL:Z98598; PIDN:CAB11240.1; GSPDB:GN00066; SPDB:SPAC13.14
 A:Experimental source: strain 972h; cosmid clB3
 C:Genetics:
 A:Gene: SPAC13.14
 A:Map position: 1
 A:Introns: 6/2; 8/2; 52/2; 160/3
 C:Superfamily: vacuolar H+-transporting ATPase 16k chain
 C:Keywords: ATP; hydrogen ion transport; hydrolase; membrane-associated complex; tran

Query Match 1.6%; Score 8; DB 2; Length 161;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 282 GAGAGAG 289
 DB 99 GAGAGAG 106
 |||||

RESULT 40
 A40814
 H+-exporting ATPase (EC 3.6.3.6) proteolipid chain, vacuolar - oat
 C:Species: *Avena sativa* (oat)
 C>Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 03-Jun-2002
 C:Accession: A40814
 R:Lal, S.; Watson, J.C.; Hansen, J.N.; Size, H.
 J. Biol. Chem. 266, 16078-16084, 1991
 A:Title: Molecular cloning and sequencing of cDNAs encoding the proteolipid subunit o
 A:Reference number: A40814; MUID:9140758; PMID:1831453
 A:Accession: A40814
 A:Molecule type: mRNA
 A:Residues: 1-165 <LAT>
 A:Cross-references: GB:M73232; NID:q166548; PIDN:AAA32712.1; PID:q166549
 C:Superfamily: vacuolar H+-transporting ATPase 16k chain
 C:Keywords: ATP; hydrogen ion transport; hydrolase; transmembrane protein

Query Match 1.6%; Score 8; DB 1; Length 165;
 Best Local Similarity 100.0%; Pred. No. 19;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 282 GAGAGAG 289
 DB 103 GAGAGAG 110
 |||||

RESULT 41
 G87307
 hypothetical protein CC0472 [imported] - *Caulobacter crescentus*
 C:Species: *Caulobacter crescentus*
 C>Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 17-May-2002
 C:Accession: G87307
 R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
 A:Title: Complete Genome Sequence of *Caulobacter crescentus*.
 A:Reference number: A87249; MUID:21173698; PMID:11259647
 A:Accession: G87307
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-178 <STO>
 A:Cross-references: GB:AE005673; NID:q13421647; PIDN:AAK22459.1; GSPDB:GN00148
 C:Superfamily: ubiquinol-cytochrome c reductase iron-sulfur protein; Rieske [2Fe-2S]

Query Match 1.6%; Score 8; DB 2; Length 178;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

A:Cross-references: GB:X02632; NID:g5686; PIDN:CAA26479.1; PID:g5687
 R:Amos, R.; Puijms, W.; Moller, W.
 FEBS Lett. 104, 85-89, 1979
 A:Title: The primary structure of ribosomal protein eL12/eL12-P from *Artemia salina* 80 S
 A:Reference number: A02775; MUID:80004136; PMID:477981
 A:Accession: A02775
 A:Molecule type: protein
 A:Residues: 1-12, 'T', 14-18, 'TX', 21-33, 'X', 35-79, 'A', 81-111 <AMO>
 A>Note: Ser-98 is partially phosphorylated; this results in two slightly different forms
 C:Superfamily: rat acidic ribosomal protein P1
 C:Keywords: phosphoprotein; protein biosynthesis; ribosome
 F:1-111/Product: ribosomal protein L12el #status experimental <MAT>
 F:98/Binding site: phosphate (Ser) (covalent) (partial) #status experimental

Query Match 1.6%; Score 8; DB 1; Length 111;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAAAGGAA 301
 |||||
 Db 72 AAAAGGAA 79

RESULT 33
 S41866
 A:acidic ribosomal protein P2 - fungus (*Cladosporium herbarum*)
 C:Species: *Cladosporium herbarum*
 C:Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 13-Aug-1999
 C:Accession: S41866
 R:Zhang, L.
 submitted to the EMBL Data Library, January 1994
 A:Description: A cDNA clone coding for a novel allergen, Cla hIII, of *Cladosporium herbarum*
 A:Reference number: S41866
 A:Accession: S41866
 A:Molecule type: mRNA
 A:Residues: 1-111 <ZHA>
 A:Cross-references: EMBL:X77253; NID:g452605; PIDN:CAA54470.1; PID:g452606
 C:Superfamily: rat acidic ribosomal protein P1
 C:Keywords: protein biosynthesis; ribosome
 F:1-111/Product: acidic ribosomal protein P2.e.B #status predicted <MAT>
 F:70-89/Region: alanine-rich
 F:99-111/Region: aspartic acid/glutamic acid-rich

Query Match 1.6%; Score 8; DB 2; Length 111;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGGA 300
 |||||
 Db 78 GAAAGGA 85

RESULT 34
 R0072B
 A:acidic ribosomal protein P2-B - *Trypanosoma cruzi*
 C:Species: *Trypanosoma cruzi*
 C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 22-Jun-1999
 C:Accession: S22950; S19948
 R:Vazquez, M.P.; Schijman, A.G.; Panabera, A.; Levin, M.J.
 Nucleic Acids Res. 20, 2893, 1992
 A:Title: A nucleotide sequence of a cDNA encoding another *Trypanosoma cruzi* acidic ribosomal
 A:Reference number: S22950; MUID:92310999; PMID:1614880
 A:Accession: S22950
 A:Molecule type: mRNA
 A:Residues: 1-112 <VAZ>
 A:Cross-references: EMBL:X65065; NID:g10633; PIDN:CAA46198.1; PID:g10634
 A>Note: it is uncertain whether Met-1 or Met-3 is the initiator
 C:Superfamily: rat acidic ribosomal protein P1
 C:Keywords: phosphoprotein; protein biosynthesis; ribosome

Query Match 1.6%; Score 8; DB 1; Length 112;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAAAGGAA 301
 |||||
 Db 81 AAAAGGAA 88

RESULT 35
 G70354
 A:hypothetical protein aq_615 - *Aquifex aeolicus*
 C:Species: *Aquifex aeolicus*
 C:Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 05-Nov-1999
 C:Accession: G70354
 R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; V.
 Nature 392, 353-358, 1998
 A:Title: The complete genome of the hyperthermophilic bacterium *Aquifex aeolicus*.
 A:Reference number: A70300; MUID:98196666; PMID:9537320
 A:Accession: G70354
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-132 <AOF>
 A:Cross-references: GB:AE000698; NID:g2983224; PIDN:AAC06830.1; PID:g2983233; GB:AE00
 A:Experimental source: strain VF5
 C:Genetics:
 A:Gene: aq_615
 C:Superfamily: *Aquifex aeolicus* hypothetical protein aq_615

Query Match 1.6%; Score 8; DB 2; Length 132;
 Best Local Similarity 100.0%; Pred. No. 16;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 289 GAAVGA 296
 |||||
 Db 35 GAAVGA 42

RESULT 36
 A83149
 A:hypothetical protein PA3971 [imported] - *Pseudomonas aeruginosa* (strain PA01)
 C:Species: *Pseudomonas aeruginosa*
 C:Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 01-Mar-2002
 C:Accession: A83149
 R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; L.
 ; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A:Title: Complete genome sequence of *Pseudomonas aeruginosa* PA01, an opportunistic pa
 A:Reference number: A82950; MUID:20437337; PMID:10984043
 A:Accession: A83149
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-143 <STO>
 A:Cross-references: GB:AE004815; GB:AE004091; NID:g9950159; PIDN:AA007358.1; GSPDB:GN
 A:Experimental source: strain PA01
 C:Genetics:
 A:Gene: PA3971
 C:Superfamily: hypothetical protein bl396

Query Match 1.6%; Score 8; DB 2; Length 143;
 Best Local Similarity 100.0%; Pred. No. 17;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 300 AAGAAAT 307
 |||||
 Db 65 AAGAAAT 72

RESULT 37
 H75262
 A:hypothetical protein - *Deinococcus radiodurans* (strain R1)
 C:Species: *Deinococcus radiodurans*
 C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 31-Mar-2000
 C:Accession: H75262

OY 290 AAVGAAA 297
 |||||||
 Db 63 AAVGAAA 70

RESULT 29

T52147
 ribosomal protein rplal [imported] - yeast (Candida albicans)

C:Species: Candida albicans
 C:Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 03-Nov-2000

C:Accession: T52147

R:Taylor, K.; Harris, D.
 submitted to the EMBL Data Library, November 1998

A:Reference number: Z25985

A:Accession: T52147

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-106 <TAV>

A:Cross-references: EMBL:AL033497; PIDN:CAA21967.1

A:Experimental source: strain 1161; cosmid CA49C10

C:Genetics:

A:Gene position: 1

A:Superfamily: rat acidic ribosomal protein P1

C:Keywords: protein biosynthesis; ribosome

Query Match 1.6%; Score 8; DB 2; Length 106;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 8; Conservative 0; Indels 0; Gaps 0;

OY 294 AAAAGGAA 301
 |||||||

Db 67 AAAAGGAA 74

RESULT 30

R6BY22
 60s acidic ribosomal protein p2.2 - fission yeast (Schizosaccharomyces pombe)

N:Alternate names: acidic ribosomal protein 2; ribosomal protein SP-140C; ribosomal prot

C:Species: Schizosaccharomyces pombe

C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 10-Dec-1999

C:Accession: B34715; A27304; S10054; T40800

R:Beltrame, M.; Bianchi, M.E.
 Mol. Cell. Biol. 10, 2341-2348, 1990

A:Title: A gene family for acidic ribosomal proteins in Schizosaccharomyces pombe: two e

A:Reference number: A54715; MUID:90220620; PMID:2325655

A:Accession: B34715

A:Molecule type: DNA

A:Residues: 1-110 <BEL>

A:Cross-references: GB:M3138; NID:g173465; PIDN:AAA35335.1; PID:g173466

R:Beltrame, M.; Bianchi, M.E.
 Nucleic Acids Res. 15, 9089, 1987

A:Title: Sequence of the cDNA for one acidic ribosomal protein of Schizosaccharomyces po

A:Reference number: A27304; MUID:88067727; PMID:3684587

A:Accession: A27304

A:Molecule type: mRNA

A:Residues: 1-110 <BEL2>

A:Cross-references: GB:Y00466; NID:g5064; PIDN:CAA6528.1; PID:g5065

A:Experimental source: strain 972h; clone p1 p887
 C:Genetics:
 A:Gene: rpa2
 A:Map position: 2
 A:Introns: 103/3
 C:Superfamily: rat acidic ribosomal protein P1
 C:Keywords: phosphoprotein; protein biosynthesis; ribosome
 F:1-110/Product: acidic ribosomal protein P2.2 #status experimental <MAT>

Query Match 1.6%; Score 8; DB 1; Length 110;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAAAGGAA 301
 |||||||

Db 76 AAAAGGAA 83

RESULT 31

R6BY24
 60s acidic ribosomal protein p2-beta - fission yeast (Schizosaccharomyces pombe)

N:Alternate names: acidic ribosomal protein 4; ribosomal protein Y112el

C:Species: Schizosaccharomyces pombe

C:Date: 31-Dec-1992 #sequence_revision 31-Dec-1992 #text_change 10-Dec-1999

C:Accession: D34715; T39961

R:Beltrame, M.; Bianchi, M.E.
 Mol. Cell. Biol. 10, 2341-2348, 1990

A:Title: A gene family for acidic ribosomal proteins in Schizosaccharomyces pombe: tw

A:Reference number: A54715; MUID:90220620; PMID:2325655

A:Accession: D34715

A:Molecule type: DNA

A:Residues: 1-110 <BEL>

A:Cross-references: EMBL:M3142; NID:g173465; PIDN:AAA35337.1; PID:g173470

R:Xiang, Z.; Aves, S.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
 submitted to the EMBL Data Library, January 1999

A:Reference number: Z21893

A:Accession: T39961

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-110 <RIA>

A:Cross-references: EMBL:AL035065; PIDN:CAA22631.1; GSPDB:GND0067; SPDB:SPBC2367.15c

A:Experimental source: strain 972h; cosmid c2367

C:Genetics:

A:Gene: rpa4

A:Map position: 2

A:Introns: 103/3

C:Superfamily: rat acidic ribosomal protein P1

C:Keywords: phosphoprotein; protein biosynthesis; ribosome

Query Match 1.6%; Score 8; DB 1; Length 110;

Best Local Similarity 100.0%; Pred. No. 13;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 296 AAGGAGA 303
 |||||||

Db 75 AAGGAGA 82

RESULT 32

R: Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998

A: Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A: Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome

A: Reference number: A70500; MUID: 98295987; PMID: 9634230

A: Accession: G70668

A: Status: nucleic acid sequence not shown; translation not shown

A: Molecule type: DNA

A: Residues: 1-1616 <COL>

A: Cross-references: GB: Z83858; GB: AL123456; NID: g3261675; PIDN: CAB06103.1; PID: g1781167

A: Experimental source: strain H37Rv

R: Smith, D.R.; Robison, K.

A: Submitted to the EMBL Data Library, September 1994

A: Description: *Mycobacterium tuberculosis* cosmid tbcc2.

A: Reference number: S73053

A: Accession: S73075

A: Molecule type: DNA

A: Residues: 'MNGARMPVRCSTEL', 48-917, 'RSEARSRCGSGRAAPPAPIGPIGP', 'PMPPIIPVEPAGFVN

A: Cross-references: EMBL: U00024; NID: g560506; PIDN: AAA50928.1; PID: g560507

C: Genetics:

A: Gene: PKS1

C: Superfamily: *Mycobacterium tuberculosis* probable polyketide synthase pks1; acyl carrier ology

C: Keywords: carrier protein

F: 84-362/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>

F: 1228-1410/Domain: short-chain alcohol dehydrogenase homology <SADH>

F: 1514-1585/Domain: acyl carrier protein homology <ACPI>

Query Match 1.8%; Score 9; DB 2; Length 1616;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 281 AGLAGLAG 289
|||||

Db 58 AGLAGLAG 66

RESULT 25

T26216

hypothetical protein W06A7.3c - *Caenorhabditis elegans*

C: Species: *Caenorhabditis elegans*

C: Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C: Accession: T26216

R: Ainscough, R.

A: Submitted to the EMBL Data Library, August 1996

A: Reference number: Z20173

A: Accession: T26216

A: Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: DNA

A: Residues: 1-2484 <NLL>

A: Cross-references: EMBL: Z78066; PIDN: CAB51467.1; GSPDB: GN00023; CESP: W06A7.3c

A: Experimental source: clone W06A7

C: Genetics:

A: Gene: CESP: W06A7.3c

A: Map position: 5

A: Introns: 1827/1; 1866/3; 1963/3; 1990/3; 2262/1; 2289/1; 2339/2; 2463/2

Query Match 1.8%; Score 9; DB 2; Length 2484;
Best Local Similarity 100.0%; Pred. No. 25;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 289 GAAGCAAA 297
|||||

Db 2207 GAAGCAAA 2215

RESULT 26

T26215

hypothetical protein W06A7.3a - *Caenorhabditis elegans*

C: Species: *Caenorhabditis elegans*

C: Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 15-Oct-1999

C: Accession: T26215

R: Ainscough, R.

A: Submitted to the EMBL Data Library, August 1996

A: Reference number: Z20173

A: Accession: T26215

A: Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: DNA

A: Residues: 1-2607 <NLL>

A: Cross-references: EMBL: Z78066; PIDN: CAB01522.2; GSPDB: GN00023; CESP: W06A7.3a

A: Experimental source: clone W06A7

C: Genetics:

A: Gene: CESP: W06A7.3a

A: Map position: 5

A: Introns: 1827/1; 1866/3; 1963/3; 1990/3; 2262/1; 2289/1; 2412/1; 2462/2; 2586/2

Query Match 1.8%; Score 9; DB 2; Length 2607;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 289 GAAGCAAA 297
|||||

Db 2207 GAAGCAAA 2215

RESULT 27

T31328

fibroin - Chinese oak silkworm

C: Species: *Antheraea pernyi* (Chinese oak silkworm)

C: Date: 22-Oct-1999 #sequence_revision 22-Oct-1999 #text_change 22-Oct-1999

C: Accession: T31328

R: Sezutsu, H.; Tamura, T.; Yukuhira, K.

A: Submitted to the EMBL Data Library, August 1998

A: Description: Characterization of the full length fibroin gene of a wild silkworm, A

A: Reference number: Z20995

A: Accession: T31328

A: Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: DNA

A: Residues: 1-2639 <SEZ>

A: Cross-references: EMBL: AF083334; NID: g3450882; PID: g3450883; PIDN: AAC32606.1

C: Genetics:

A: Introns: 14/3

Query Match 1.8%; Score 9; DB 2; Length 2639;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 297 AGGAGAAA 305
|||||

Db 117 AGGAGAAA 125

RESULT 28

T35731

hypothetical protein SC7H1.36c SC7H1.36c - *Streptomyces coelicolor* (fragment)

C: Species: *Streptomyces coelicolor*

C: Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 05-Nov-1999

C: Accession: T35731

R: Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

A: Submitted to the EMBL Data Library, January 1998

A: Reference number: Z21548

A: Accession: T35731

A: Status: preliminary; translated from GB/EMBL/DBJ

A: Molecule type: DNA

A: Residues: 1-81 <MRP>

A: Cross-references: EMBL: AL021411; PIDN: CA116223.1; GSPDB: GN00070; SCOPDB: SC7H1.36c

A: Experimental source: strain A3(2)

C: Genetics:

A: Gene: SCOPDB: SC7H1.36c

Query Match 1.6%; Score 8; DB 2; Length 81;
Best Local Similarity 100.0%; Pred. No. 10;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

F:28-860/Product: elastin #status predicted <MAT>
F:850-855/Disulfide bonds: #status predicted

Query Match 1.8%; Score 9; DB 1; Length 860;
Best Local Similarity 100.0%; Pred. No. 9.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 298 GGAAGAAA 306
|||||
DB 113 GGAAGAAA 121

RESULT 20

140889 sarcosine oxidase (EC 1.5.3.1) alpha chain [validated] - Corynebacterium sp.

C:Species: Corynebacterium sp.

C:Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 26-May-2000

C:Accession: 140889

R:Chlumsky, L.J.; Zhang, L.; Jorns, M.S.

A:Title: Sequence analysis of sarcosine oxidase and nearby genes reveals homologues with

A:Accession: 140889

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-967 <RES>

A:Cross-references: EMBL:U23955; NID:927587; PIDN:AAC43461.1; PID:927591

C:Genetics:

A:Gene: socA

C:Complex: heterotetramer: alpha (PIR:140889), beta (PIR:140887), gamma (PIR:140888), an

C:Function: EC 1.5.3.1 [validated, MUID:95355441]; catalyzes the oxidative demethylat

A:Description: EC 1.5.3.1 [validated, MUID:95355441]; FAD plus 1 mol FAD covalently attac

A:Note: enzyme contains 1 mol of noncovalently bound FAD plus 1 mol FAD covalently attac

C:Keywords: oxidoreductase

Query Match 1.8%; Score 9; DB 2; Length 967;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 289 GAAVGAAGA 297
|||||
DB 434 GAAVGAAGA 442

RESULT 21

113807 potassium channel protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 13-Aug-1999 #sequence_revision 13-Aug-1999 #text_change 17-Nov-2000

C:Accession: 113807

R:Goldstein, S.A.; Price, L.A.; Rosenthal, D.N.; Pausch, M.H.

A:Title: ORK1, a potassium-selective leak channel with two pore domains cloned from Dros

A:Reference number: 217770; MUID:97075152; PMID:8917578

A:Accession: 113807

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1001 <GOL>

A:Cross-references: EMBL:U55321; NID:93808067; PIDN:AAC69250.1

C:Genetics:

A:Gene: ORK1

A:Cross-references: FlyBase:FBgn0017561

Query Match 1.8%; Score 9; DB 2; Length 1001;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAAAGGAAG 302
|||||
DB 764 AAAAGGAAG 772

RESULT 22

B70985 probable polyketide synthase - Mycobacterium tuberculosis (strain H37Rv)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 26-May-2000

C:Accession: B70985

R:Cole, S.T.; Broesch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holtroyd,

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987; PMID:9634230

A:Accession: B70985

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-1017 <COL>

A:Cross-references: GB:295617; GB:AL123456; NID:93242249; PIDN:CAB09100.1; PID:e31726

A:Experimental source: strain H37Rv

C:Genetics:

A:Gene: pks9

C:Superfamily: [acyl-carrier-protein] S-malonyltransferase homology; 3-oxoacyl-[acyl-

C:Keywords: carrier protein

F:24-392/Domain: 3-oxoacyl-[acyl-carrier-protein] synthase I homology <OAS2>

F:504-782/Domain: [acyl-carrier-protein] S-malonyltransferase homology <AMT1>

F:864-942/Domain: acyl carrier protein homology <ACP2>

Query Match 1.8%; Score 9; DB 2; Length 1017;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 281 AGIAGIAG 289
|||||
DB 478 AGIAGIAG 486

RESULT 23

T00117 dve protein - fruit fly (Drosophila melanogaster)

C:Species: Drosophila melanogaster

C:Date: 22-Jan-1999 #sequence_revision 22-Jan-1999 #text_change 17-Nov-2000

C:Accession: T00117

R:Nakagoshi, H.; Hoshi, M.; Nabeshima, Y.; Matsuzaki, F.

Genes Dev. 12, 2724-2734, 1998

A:Title: A novel homeobox gene mediates the DPP signal to establish functional specif

A:Reference number: Z14112; MUID:98402482; PMID:9732270

A:Accession: T00117

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-1019 <NAK>

A:Cross-references: EMBL:AB010299; NID:93523072; PIDN:BA032660.1; PID:93523073

C:Genetics:

A:Gene: dve

A:Cross-references: FlyBase:FBgn0020307

A:Map position: 2

Query Match 1.8%; Score 9; DB 2; Length 1019;

Best Local Similarity 100.0%; Pred. No. 11;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAAAGGAAG 302
|||||
DB 477 AAAAGGAAG 485

RESULT 24

G70668 polyketide synthase pks1 - Mycobacterium tuberculosis (strain H37Rv)

N:Alternate names: polyketide synthase pks002c

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 20-Jun-2000

C:Accession: G70668; S73075

A:Experimental source: strain Nigg (MoPn)
 C:Genetics:
 A:Gene: TC0867
 C:Superfamily: conserved hypothetical protein TC0867

Query Match 1.8%; Score 9; DB 2; Length 491;
 Best Local Similarity 100.0%; Pred. No. 5.9;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 447 FWOQASKIA 455
 |||||||
 Db 447 FWOQASKIA 455

RESULT 17

EABO
 N:Alternate names: splice form a - bovine
 N:Contains: elastin precursor, splice form b; elastin precursor, splice form c
 C:Species: Bos primigenius taurus (cattle)
 C:Date: 08-Jun-1989 #sequence_revision 26-Jul-1996 #text_change 22-Jun-1999
 C:Accession: A31865; A26728; B26728; C26728; A22343; I45886
 R:Ieh, H.; Anderson, N.; Ornstein-Goldstein, N.; Bashir, M.M.; Rosenbloom, J.C.; Abrams, Biochemistry 28, 2365-2370, 1989
 A:Title: Structure of the bovine elastin gene and S1 nuclease analysis of alternative splicing
 A:Reference number: A31865; MUID:89274159; PMID:2543440
 A:Accession: A31865
 A:Molecule type: DNA
 A:Residues: 1-27 <YES>
 A:Cross-references: GB:J02855; NID:g340504; PIDN:AAA30776.1; PID:g552339
 J:Raju, K.; Anwar, R.A.
 J. Biol. Chem. 262, 5755-5762, 1987
 A:Title: Primary structures of bovine elastin a, b, and c deduced from the sequences of cDNAs
 A:Reference number: A92640; MUID:87194772; PMID:3032943
 A:Accession: A26728
 A:Molecule type: mRNA
 A:Residues: 1,'RS','4-11','E','13-636','V','638-747 <RAJ>
 A:Cross-references: GB:J02717; NID:g163019; PIDN:AAA30503.1; PID:g163020
 A:Accession: B26728
 A:Molecule type: mRNA
 A:Residues: 1,'RS','4-11','E','13-225,240-636','V','638-747 <RAJ>
 A:Cross-references: GB:K03505; NID:g163025; PIDN:AAA30505.1; PID:g163026
 A:Accession: C26728
 A:Molecule type: mRNA
 A:Residues: 1,'RS','4-11','E','13-225,260-636','V','638-747 <RAJ>
 A:Cross-references: GB:K03506; NID:g163027; PIDN:AAA30506.1; PID:g163028
 R:Ciella, G.; May, M.; Ornstein-Goldstein, N.; Indik, Z.; Morrow, S.; Ieh, H.S.; Rosenb Biochemistry 24, 3075-3080, 1985
 A:Title: Structure of the 3' portion of the bovine elastin gene.
 A:Reference number: A22343; MUID:85280426; PMID:2992576
 A:Accession: A22343
 A:Molecule type: DNA
 A:Residues: 613-747 <CIC>
 A:Cross-references: GB:M20415
 R:Rosenbloom, J.
 Lab. Invest. 51, 605-623, 1984
 A:Title: Biology of disease: Elastin: Relation of protein and gene structure to disease.
 A:Reference number: I45885; MUID:8505254; PMID:6150137
 A:Accession: I45886
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 678-683,685-747 <ROS>
 A:Cross-references: GB:M31898; NID:g163015; PIDN:AAA96417.1; PID:g163018
 R:Brown, P.L.; Mechem, L.; Tisdale, C.; Mechem, R.P.
 Biochem. Biophys. Res. Commun. 186, 549-555, 1992
 A:Title: The cysteine residues in the carboxy terminal domain of tropoelastin form an intramolecular disulfide bond
 A:Contents: annotation, disulfide bonds
 C:Comment: The term tropoelastin refers to a soluble precursor form of the extracellular line oxidase activity.
 C:Genetics:
 A:introns: 634/3; 653/3; 676/3; 689/3; 707/3; 716/3; 733/3
 A:Note: the list of introns is incomplete

C:Superfamily: elastin
 C:Keywords: alternative splicing; extracellular matrix; glycoprotein; hydroxyllysine
 F:1-747/Product: elastin precursor, splice form a #status predicted <EPA>
 F:1-225,260-747/Product: elastin precursor, splice form c #status predicted <EPC>
 F:1-225,240-747/Product: elastin precursor, splice form b #status predicted <EPB>
 F:1-26/DNA: signal sequence #status predicted <SIG>
 F:27-747/Product: elastin #status predicted <MAT>
 F:105,109,252,271,275,324,327,400,404,407,445,448,489,493,544,548,552,606,609,645,649
 F:737-742/Disulfide bonds: #status experimental

Query Match 1.8%; Score 9; DB 1; Length 747;
 Best Local Similarity 100.0%; Pred. No. 8.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 298 GGAAGAAA 306
 |||||||
 Db 95 GGAAGAAA 103

RESULT 18

S59623
 tropoelastin - sheep
 C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
 C:Date: 23-Aug-1996 #sequence_revision 13-Mar-1997 #text_change 24-Oct-1997
 C:Accession: S59623; A24758
 R:Mauch, J.C.; Sandberg, L.B.; Roos, P.J.; Jimenez, F.; Cristiano, A.M.; Deak, S.B.; Matrix Biol. 14, 635-641, 1994
 A:Title: Extensive alternate exon usage at the 5' end of the sheep tropoelastin gene.
 A:Reference number: S59623
 A:Accession: S59623
 A:Status: preliminary; not compared with conceptual translation
 A:Molecule type: mRNA
 A:Residues: 1-770 <MAU>
 R:Iyon, K.; Davidson, J.M.; Boyd, C.; May, M.; Luvall, P.; Ornstein-Goldstein, N.; S Arch. Biochem. Biophys. 241, 684-691, 1985
 A:Title: Analysis of the 3' region of the sheep elastin gene.
 A:Reference number: A24758; MUID:85305763; PMID:3839997
 A:Accession: A24758
 A:Molecule type: mRNA
 A:Residues: 655-669,671-716,732-770 <YOO>
 C:Superfamily: elastin
 C:Keywords: alternative splicing; extracellular matrix; glycoprotein; hydroxyllysine
 F:760-765/Disulfide bonds: #status predicted

Query Match 1.8%; Score 9; DB 2; Length 770;
 Best Local Similarity 100.0%; Pred. No. 8.8;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 298 GGAAGAAA 306
 |||||||
 Db 97 GGAAGAAA 105

RESULT 19

EAMS
 elastin precursor - mouse
 N:Alternate names: tropoelastin
 C:Species: Mus musculus (house mouse)
 C:Date: 18-Aug-1995 #sequence_revision 16-Aug-1996 #text_change 22-Jun-1999
 C:Accession: A55721
 R:Wyner, K.S.; Sechler, J.L.; Boyd, C.D.; Passmore, H.C.
 Genomics 23, 125-131, 1994
 A:Title: Use of an intron length polymorphism to localize the tropoelastin gene to mouse chromosome 10
 A:Reference number: A55721; MUID:95130069; PMID:7829060
 A:Accession: A55721
 A:Molecule type: mRNA
 A:Residues: 1-860 <WY>
 A:Cross-references: GB:U08210; NID:g473273; PIDN:AAA80155.1; PID:g473274
 C:Genetics:
 A:Map position: 5
 C:Superfamily: elastin
 C:Keywords: alternative splicing; extracellular matrix; glycoprotein; hydroxyllysine
 F:1-27/DNA: signal sequence #status predicted <SIG>

C:Superfamily: peripheral-type benzodiazepine receptor

Query Match 1.8%; Score 9; DB 2; Length 169;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 302 GAAATTTVA 310
|||||
DB 117 GAAATTTVA 125

RESULT 12

T34870
probable membrane protein - Streptomyces coelicolor

C:Species: Streptomyces coelicolor

C>Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 08-Sep-2000

C:Accession: T34870

R:Seeger, K.J.; Harris, D.; James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, August 1999

A:Reference number: Z21560

A:Accession: T34870

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-243 <SEP>

A:Cross-references: EMBL:AL109849; PIDN:CAB52862.1; GSPDB:GN00070; SCOEDB:SC3A3.05C

A:Experimental source: strain A3(2)

C:Genetics:

A:Gene: SCOEDB:SC3A3.05C

C:Superfamily: Streptomyces coelicolor probable membrane protein SC3A3.05C

Query Match 1.8%; Score 9; DB 2; Length 243;
Best Local Similarity 100.0%; Pred. No. 3.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 282 GAGLAAGA 290
|||||
DB 63 GAGLAAGA 71

RESULT 13

AE2972
secretion protein, HlyD family [imported] - Agrobacterium tumefaciens (strain C58, Dupont)

C:Species: Agrobacterium tumefaciens

C>Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 11-Jan-2002

C:Accession: AE2972

R:Wood, D.W.; Setubal, J.C.; Kaut, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moo, I.

erage, G.; Giller, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell

; Karp, P.; Romero, P.; Zhang, S.

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,

ster, E.W.

A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: AE2972

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-437 <KUR>

A:Cross-references: GB:AE008689; PIDN:AAI44195.1; PID:g17741773; GSPDB:GN00187

A:Experimental source: strain C58 (Dupont)

C:Genetics:

A:Gene: prfE

A:Map position: linear chromosome

Query Match 1.8%; Score 9; DB 2; Length 437;
Best Local Similarity 100.0%; Pred. No. 5.3;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 283 LAGLAAGA 291
|||||
DB 19 LAGLAAGA 27

RESULT 14

P98310

rhizobium secretion protein rseP (AF141932) [imported] - Agrobacterium tumefaciens

C:Species: Agrobacterium tumefaciens

C>Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 11-Jan-2002

C:Accession: P98310

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Gold

A.; Liu, F.; Wollam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz,

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium

A:Reference number: A97359; PMID:11743194

A:Accession: P98310

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-474 <KUR>

A:Cross-references: GB:AE007870; PIDN:AAK90008.1; PID:g15159979; GSPDB:GN00170

C:Genetics:

A:Gene: AGR_L_2881

A:Map position: linear chromosome

Query Match 1.8%; Score 9; DB 2; Length 474;
Best Local Similarity 100.0%; Pred. No. 5.7;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 283 LAGLAAGA 291
|||||
DB 56 LAGLAAGA 64

RESULT 15

D71497
hypothetical protein CT578 - Chlamydia trachomatis (serotype D, strain UW3/Cx)

C:Species: Chlamydia trachomatis

C>Date: 13-Sep-1998 #sequence_revision 13-Sep-1998 #text_change 18-Aug-2000

C:Accession: D71497

R:Stephens, R.S.; Kalman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitche

Science 282, 754-759, 1998

A:Title: Genome sequence of an obligate intracellular pathogen of humans: Chlamydia t

A:Reference number: A71570; MUID:9900809; PMID:9784136

A:Accession: D71497

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-487 <ARN>

A:Cross-references: GB:AE001328; GB:AE001273; NID:g3329015; PIDN:AC68180.1; PID:g332

A:Experimental source: serotype D, strain UW-3/Cx

C:Genetics:

A:Gene: CT578

C:Superfamily: conserved hypothetical protein TC0867

Query Match 1.8%; Score 9; DB 2; Length 487;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 447 FMOQASKIA 455
|||||
DB 443 FMOQASKIA 451

RESULT 16
F81655
conserved hypothetical protein TC0867 [imported] - Chlamydia muridarum (strain Nig9)

C:Species: Chlamydia muridarum, Chlamydia trachomatis MoPn

C>Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 18-Aug-2000

C:Accession: F81655

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke

, C.; Dodson, R.; Gwinn, M.; Nelson, W.; Deboy, R.; Kolonay, J.; McClarty, G.; Salzbe

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis MoPn and Chlamydia pneumoniae AR39

A:Reference number: A81500; MUID:20150255; PMID:10684935

A:Accession: F81655

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-491 <REV>

A:Cross-references: GB:AE002353; GB:AE002160; NID:g7190891; PIDN:AAF39663.1; PID:g719

A:Residues: 1-111 <ACH>
 A:Cross-references: EMBL:X78223; NID:g467626; PIDN:CAA5067.1; PID:g467627
 C:Superfamily: rat acidic ribosomal protein p1

Query Match 1.8%; Score 9; DB 2; Length 111;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGGAA 301
 DB 77 GAAAGGAA 85

RESULT 7

T49895
 oleosin-like protein - Arabidopsis thaliana

M:Alternate names: protein T211.270

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 17-Nov-2000

C:Accession: T49895

R:Bevan, M.; Murphy, G.; Riddle, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Rudd, S.; Le
 Submitted to the Protein Sequence Database, April 2000

A:Reference number: Z24493

A:Accession: T49895

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-153 <BEV>

A:Cross-references: EMBL:AL163912; GSPDB:GN00063; ATSP:T211.270

A:Experimental source: cultivar Columbia; BAC clone T211

C:Genetics:

A:Gene: ATSP:T211.270

A:Map position: 5

A:Introns: 106/2

C:Superfamily: oleosin

Query Match 1.8%; Score 9; DB 2; Length 153;
 Best Local Similarity 100.0%; Pred. No. 2.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGGAA 301
 DB 119 GAAAGGAA 127

RESULT 8

I38724
 mitochondrial benzodiazepine receptor - human

C:Species: Homo sapiens (man)

C:Date: 29-May-1998 #sequence_revision 29-May-1998 #text_change 29-Sep-1999

C:Accession: I38724

R:Yakovlev, A.G.; Ruffo, M.; Jurka, J.; Krueger, K.E.

Gene 155, 201-205, 1995

A:Title: Comparison of repetitive elements in the third intron of human and rodent mitoc

A:Reference number: I38724; MUID:95237610; PMID:7721091

A:Accession: I38724

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-169 <RES>

A:Cross-references: EMBL:U12421; NID:g529945; PIDN:AAA83252.1; PID:g529946

C:Genetics:

A:Gene: MBR

A:Introns: 61/2; 107/3

C:Superfamily: peripheral-type benzodiazepine receptor

C:Keywords: mitochondrion

Query Match 1.8%; Score 9; DB 2; Length 169;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 302 GAAATTVA 310
 DB 117 GAAATTVA 125

RESULT 9

S14257
 benzodiazepine receptor, peripheral-type - human

C:Species: Homo sapiens (man)

C:Date: 18-Feb-1994 #sequence_revision 10-Nov-1995 #text_change 29-Sep-1999

C:Accession: S14257

R:Riond, J.; Mattei, M.G.; Kaghad, M.; Dumont, X.; Guillemot, J.C.; Le Fur, G.; Caput

Eur. J. Biochem. 195, 305-311, 1991

A:Title: Molecular cloning and chromosomal localization of a human peripheral-type be

A:Reference number: S14257; MUID:91146565; PMID:1847678

A:Accession: S14257

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-169 <RTO>

A:Cross-references: GB:M36035; NID:g184333; PIDN:AAA03652.1; PID:g306883

C:Genetics:

A:Gene: GDB:B2RP

A:Cross-references: GDB:127347; OMIM:109610

A:Map position: 22q13.31-22qter

C:Superfamily: peripheral-type benzodiazepine receptor

C:Keywords: mitochondrion; transmembrane protein

Query Match 1.8%; Score 9; DB 2; Length 169;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 302 GAAATTVA 310
 DB 117 GAAATTVA 125

RESULT 10
 I38105
 peripheral benzodiazepine receptor - human

C:Species: Homo sapiens (man)

C:Date: 24-May-1996 #sequence_revision 24-May-1996 #text_change 29-Sep-1999

C:Accession: I38105

R:Lin, D.; Chang, Y.J.; Strauss, J.F.; Miller, W.L.

Genomics 18, 643-650, 1993

A:Title: The human peripheral benzodiazepine receptor gene: cloning and characterizat

A:Reference number: A49361; MUID:94140364; PMID:8307574

A:Accession: I38105

A:Status: preliminary; translated from GB/EMBL/DDBJ

A:Molecule type: DNA

A:Residues: 1-169 <RES>

A:Cross-references: GB:L21954; NID:g483405; PIDN:AAA18228.1; PID:g488425

C:Genetics:

A:Gene: MBR

A:Introns: 61/2; 107/3

C:Superfamily: peripheral-type benzodiazepine receptor

Query Match 1.8%; Score 9; DB 2; Length 169;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 302 GAAATTVA 310
 DB 117 GAAATTVA 125

RESULT 11
 J60149
 peripheral benzodiazepine receptor - human

C:Species: Homo sapiens (man)

C:Date: 10-Jul-1998 #sequence_revision 10-Jul-1998 #text_change 21-Nov-1998

C:Accession: J60149

R:Xu, G.H.; Hu, B.R.; Luo, C.O.; Chen, R.Z.; Qiu, P.X.

Chinese Biochem. J. 13, 423-427, 1997

A:Title: Cloning of the cDNAs encoding human peripheral benzodiazepine receptor and i

A:Reference number: J60149

A:Accession: J60149

A:Molecule type: mRNA

A:Residues: 1-169 <XUA>

Db 481 SGATAGA 487

THE

RESULT 3

C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
 C:Accession: G86591
 R:Shirai, M.; Hirkawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Ise,
 Nucleic Acids Res. 28, 2311-2314, 2000
 A:Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A:Reference number: A86491; MUID:20303049; PMID:10871362
 A:Accession: G86591
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-488 <STD>
 A:Cross-references: GB:BA000008; NID:58979183; PIDN:BA099017.1; GSPDB:GN00142
 A:Experimental source: strain J138
 C:Genetics:
 A:Gene: CPJ0809
 C:Superfamily: conserved hypothetical protein TC0867

Query Match	98.2%	Score 487;	DB 2;	Length 488;
Best Local Similarity	100.0%	Pred. No. 0;		
Matches 487;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

Qy	4	MISSSSSGPDNKNIMSOVLSTPOGVPOOKLSGNETKOIQOROCKNTMEMEDAT	63
Db	1	MISSSSSGDNKNIMSOVLSTPOGVPOOKLSGNETKOIQOROCKNTMEMEDAT	60
Qy	64	ASGKXTSTSTTETETAPOQVAAKESSESOKAGADTVSGAAATTSNTATKIAMOTSI	123
Db	61	ASGKXTSTSTTETETAPOQVAAKESSESOKAGADTVSGAAATTSNTATKIAMOTSI	120
Qy	124	EEAKSMESTLESLOSLSAAOKKEVEAVVVAALSGKSSGSKLEPELPKPGVPRSEVI	188
Db	121	EEAKSMESTLESLOSLSAAOKKEVEAVVVAALSGKSSGSKLEPELPKPGVPRSEVI	180
Qy	184	EIGLALAKAIQTLGEATKSALSNVASTOQAODTNKLCLEKOAIKIDKEREYODMKRAE	243
Db	181	EIGLALAKAIQTLGEATKSALSNVASTOQAODTNKLCLEKOAIKIDKEREYODMKRAE	240
Qy	244	OKSKOLEGMDVNYNMIASVAIVISIVAAIFPCGAGLGLAAGAAVGAAGAAGAGA	303
Db	241	OKSKOLEGMDVNYNMIASVAIVISIVAAIFPCGAGLGLAAGAAVGAAGAAGAGA	300
Qy	304	AAATVATQITVOAVVQAVKAVITAVROAITTAIKAAVSSIKAFITLYKATIAKAISK	363
Db	301	AAATVATQITVOAVVQAVKAVITAVROAITTAIKAAVSSIKAFITLYKATIAKAISK	360
Qy	364	GISKVPAKGMOMAKNPPKLSKYISLSITKWTYVGVVVVAAPALGKGIMQMOJSEMOON	423
Db	361	GISKVPAKGMOMAKNPPKLSKYISLSITKWTYVGVVVVAAPALGKGIMQMOJSEMOON	420
Qy	424	VAOFOREVGLQAAADMISMTFOWOQASKIASKOTGESNENTOKATIKLGAOILKAYAAI	483
Db	421	VAOFOREVGLQAAADMISMTFOWOQASKIASKOTGESNENTOKATIKLGAOILKAYAAI	480
Qy	484	SGAIIAGA 490	
Db	481	SGAIIAGA 487	

RESULT 4

A43742 female sterile homeotic protein, 205K - fruit fly (*Drosophila melanogaster*)
N:Alternate names: membrane protein 1sh, 205K
N:Contains: female sterile homeotic protein, 110K
C:Species: *Drosophila melanogaster*
C:Date: 03-Mar-1993 #sequence_revision 03-Mar-1993 #text_change 20-Sep-1995
C:Accession: A43742; B43742

R:Haynes, S.R.; Mozer, B.A.; Bhatia-Dey, N.; Dawid, I.B.

Dev. Biol. 134, 246-257, 1989.

A: Article; the Bioscience Resource Project
A: Reference number; PMID: 2567251
A: Reference number; MUID: 89276730; PMID: 2567251

A;Accession: A43742

```
A;status: preliminary
A;molecule type: mF
```

A;Residues: 1-2038 <HAY>

A;Accession: B43742

A;Status: preliminary

A;Molecule type: mRNA
A;Residues: 1-1106 <HA2>

A; Cross-references: EMBL

C;Genetics:
A:Gene: fsh

A; Cross-references: FLYE

C;Superfamily: unassigned
C;keywords: alternative

F;1-2038/Product: female

F;1-1106/Product: female
F;59-116/Domatin: bromod

F;503-560/Domain: bromod

Answer Match

query match
Best Local Similarity

Matches 11; Conserv

294 AAAAGGAGAA 30

[illegible]

DD 343 AAAAGGAAAGAA 3-

Query Match	2.2%;	Score 11;	DB 2;	Length 2038;
Best Local Similarity	100.0%;	Pred. No. 0.29;		
Matches 11;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

QY	294	AAAAGGAAGAA	304
Db	343	AAAAGGAAGAA	353

RESULT 5

T30714
hypothetical protein 112L - Molluscum contagiosum virus 1
N:Alternate names: MC112L
C:Species: Molluscum contagiosum virus 1
C:Date: 05-Nov-1999 #sequence_revision 05-Nov-1999 #text_change 21-Jul-2000
C:Accession: T30714
R:Seventich, T.G.; Bugert, J.J.; Sisler, J.R.; Koonin, E.V.; Darai, G.; Moss, B.
Science 273, 813-816, 1996
A:Title: Genome sequence of a human tumorigenic poxvirus: Prediction of specific host
A:Reference number: Z20876; MUID:96325459; PMID:8670425
A:Accession: T30714
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-118 <SEN>
A:Cross-references: EMBL:U60315; NID:g1491943; PIDN:AAC55240.1; PID:g1492055
C:Genetics:
A:Note: MC112L

Query Match	2.0%;	Score 10;	DB 2;	Length 128;
Best Local Similarity	100.0%;	Pred. No. 0.21;		
Matches	10;	Conservative	0;	Mismatches
			0;	Indels
				Gaps
				0;

QY	287	AAGAAVGAAA	296
Db	83	AAGAAVGAAA	92

RESULT 6

S43115
 acidic ribosomal protein p2 - fungus (Cladosporium herbarum)
 C:Species: Cladosporium herbarum
 C:Date: 19-Mar-1997 #sequence_revision 06-Jun-1997 #text_change 13-Aug-1999
 C:Accession: S43115
 R:Arbocz, G.; Oberkofler, H.; Simon, B.; Lechenauer, E.; Unger, A.; Kandler, D.; Prill
 submitted to the EMBL Data Library, March 1994
 A:Description: Molecular characterization of allergens of Cladosporium herbarum and A
 A:Reference number: S43108
 A:Accession: S43115
 A:Molecule type: mRNA


```
979 6 1.2 155 2 F84367 hypothetical prote
980 6 1.2 155 2 T23925 hypothetical prote
981 6 1.2 156 2 S41771 glycine-rich RNA-b
982 6 1.2 156 2 T470712 hypothetical prote
983 6 1.2 156 2 T47739 hypothetical prote
984 6 1.2 156 2 T16531 hypothetical prote
985 6 1.2 156 2 G72591 hypothetical prote
986 6 1.2 157 2 S49448 oleosin - rape
987 6 1.2 157 2 T33892 hypothetical prote
988 6 1.2 157 2 S13803 homeotic protein A
989 6 1.2 157 2 F87574 CBS domain protein
990 6 1.2 157 2 B69538 conserved hypotet
991 6 1.2 157 2 AE2223 hypothetical prote
992 6 1.2 158 2 H75028 probable H+-transp
993 6 1.2 158 2 G75542 probable RNA methy
994 6 1.2 158 2 C32055 nifx protein - Azot
995 6 1.2 158 2 C86937 probable membrane
996 6 1.2 158 2 E87271 hypothetical prote
997 6 1.2 158 2 H85620 hypothetical prote
998 6 1.2 158 2 E97500 conserved hypotet
999 6 1.2 158 2 A12718 type 4 pilin - Elk
1000 6 1.2 159 2 A47699
```

ALIGNMENTS

RESULT 1

F72031

c5578 hypothetical protein - Chlamydomophila pneumoniae (strain CML029)

C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C:Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 18-Aug-2000

A:Accession: F72031

R:Kalan, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;

Nature Genet. 21, 385-389, 1999

A:Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.

A:Reference number: A72000; MUID:99206606; PMID:10192388

A:Accession: F72031

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-493 <ARN>

A:Cross-references: GB:AE001662; GB:AE001363; NID:94377118; PIDN:AAD18947.1; PID:94377112

A:Experimental source: strain CML029

C:Genetics:

A:Gene: CPN0809

C:Superfamily: conserved hypothetical protein TC0867

```
Query Match          99.4%; Score 493; DB 2; Length 493;
Best local Similarity 100.0%; Pred. No. 0;
Matches 493; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 4 MSISSSSGPDNKNIMSOVLTSTPGVPOODKLSGNETKOIOOTROGKNTMESDATTIAG 63
    |||||||
DB 1 MSISSSSGPDNKNIMSOVLTSTPGVPOODKLSGNETKOIOOTROGKNTMESDATTIAG 60
QY 64 ASGKDKTSSSTTKETAPQOGVAAGKSSSESOKAGADTVGSGAATTAASNTATIAAQTST 123
    |||||||
DB 61 ASGKDKTSSSTTKETAPQOGVAAGKSSSESOKAGADTVGSGAATTAASNTATIAAQTST 120
QY 124 EASKSMESTLESLOSLSAAQMEKEVAVVVAALSGKSSGSAKLETPELPRGVTTPRSEVT 183
    |||||||
DB 121 EASKSMESTLESLOSLSAAQMEKEVAVVVAALSGKSSGSAKLETPELPRGVTTPRSEVT 180
QY 184 EIGLALAKAIQTLGEATKSLSNVASTQAQADOTNKLGLEKQAIKIDKEEREYOEKMAAE 243
    |||||||
DB 181 EIGLALAKAIQTLGEATKSLSNVASTQAQADOTNKLGLEKQAIKIDKEEREYOEKMAAE 240
QY 244 OKSKDLEGTMDTVNTVMIAVSAITVISAALFTGCGAGLAGLAAGAAGAAAGGAAGA 303
    |||||||
DB 241 OKSKDLEGTMDTVNTVMIAVSAITVISAALFTGCGAGLAGLAAGAAGAAAGGAAGA 300
QY 304 AATTVATQITVOAVVOAKQAVITVAVROAITAIRAANKSGIKAFIKTLVKAIAAISK 363
    |||||||
```

```
DB 301 AATTVATQITVOAVVOAKQAVITVAVROAITAIRAANKSGIKAFIKTLVKAIAAISK 360
QY 364 GISKVFAKGTOMIAKNPKISKVSSLSKWTVGVGVAAPALGKIMOMLSMOON 423
    |||||||
DB 361 GISKVFAKGTOMIAKNPKISKVSSLSKWTVGVGVAAPALGKIMOMLSMOON 420
QY 424 VAOFQKEVGLQAAMIMISFETPMOQASKIASKQGESNEMTOKATKLAGAOLTKAAVAAI 483
    |||||||
DB 421 VAOFQKEVGLQAAMIMISFETPMOQASKIASKQGESNEMTOKATKLAGAOLTKAAVAAI 480
QY 484 SGAIAGAHKTNMF 496
    |||||||
DB 481 SGAIAGAHKTNMF 493
```

RESULT 2

AB1507

conserved hypothetical protein CP1062 (imported) - Chlamydomophila pneumoniae (strain A

C:Species: Chlamydomophila pneumoniae, Chlamydia pneumoniae

C:Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 18-Aug-2000

A:Accession: AB1507

R:Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hicke

, C.; Dodson, R.; Gwin, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzbe

Nucleic Acids Res. 28, 1397-1406, 2000

A:Title: Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39

A:Reference number: AB1500; MUID:20150255; PMID:10684935

A:Accession: AB1507

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-488 <REA>

A:Cross-references: GB:AE002263; GB:AE002161; NID:97189971; PIDN:AAF38835.1; PID:9718

A:Experimental source: strain AR39, HL cells

C:Genetics:

A:Gene: CP1062

C:Superfamily: conserved hypothetical protein TC0867

```
Query Match          98.2%; Score 487; DB 2; Length 488;
Best local Similarity 100.0%; Pred. No. 0;
Matches 487; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 4 MSISSSSGPDNKNIMSOVLTSTPGVPOODKLSGNETKOIOOTROGKNTMESDATTIAG 63
    |||||||
DB 1 MSISSSSGPDNKNIMSOVLTSTPGVPOODKLSGNETKOIOOTROGKNTMESDATTIAG 60
QY 64 ASGKDKTSSSTTKETAPQOGVAAGKSSSESOKAGADTVGSGAATTAASNTATIAAQTST 123
    |||||||
DB 61 ASGKDKTSSSTTKETAPQOGVAAGKSSSESOKAGADTVGSGAATTAASNTATIAAQTST 120
QY 124 EASKSMESTLESLOSLSAAQMEKEVAVVVAALSGKSSGSAKLETPELPRGVTTPRSEVT 183
    |||||||
DB 121 EASKSMESTLESLOSLSAAQMEKEVAVVVAALSGKSSGSAKLETPELPRGVTTPRSEVT 180
QY 184 EIGLALAKAIQTLGEATKSLSNVASTQAQADOTNKLGLEKQAIKIDKEEREYOEKMAAE 243
    |||||||
DB 181 EIGLALAKAIQTLGEATKSLSNVASTQAQADOTNKLGLEKQAIKIDKEEREYOEKMAAE 240
QY 244 OKSKDLEGTMDTVNTVMIAVSAITVISAALFTGCGAGLAGLAAGAAGAAAGGAAGA 303
    |||||||
DB 241 OKSKDLEGTMDTVNTVMIAVSAITVISAALFTGCGAGLAGLAAGAAGAAAGGAAGA 300
QY 304 AATTVATQITVOAVVOAKQAVITVAVROAITAIRAANKSGIKAFIKTLVKAIAAISK 363
    |||||||
DB 301 AATTVATQITVOAVVOAKQAVITVAVROAITAIRAANKSGIKAFIKTLVKAIAAISK 360
QY 364 GISKVFAKGTOMIAKNPKISKVSSLSKWTVGVGVAAPALGKIMOMLSMOON 423
    |||||||
DB 361 GISKVFAKGTOMIAKNPKISKVSSLSKWTVGVGVAAPALGKIMOMLSMOON 420
QY 424 VAOFQKEVGLQAAMIMISFETPMOQASKIASKQGESNEMTOKATKLAGAOLTKAAVAAI 483
    |||||||
DB 421 VAOFQKEVGLQAAMIMISFETPMOQASKIASKQGESNEMTOKATKLAGAOLTKAAVAAI 480
QY 484 SGAIAGAHKTNMF 490
```


687	6	1.2	80	2	S62132	H+-transporting tw
688	6	1.2	80	2	S72757	Bl196_C2_181 prote
689	6	1.2	81	2	G82561	hypothetical prote
690	6	1.2	82	2	J01181	outer envelope mem
691	6	1.2	82	2	E85891	conserved hypotet
692	6	1.2	83	2	T23070	hypothetical prote
693	6	1.2	87	2	B82115	conserved hypotet
694	6	1.2	87	2	E95371	conserved hypotet
695	6	1.2	88	1	MMAD9	early E1b 9K prote
696	6	1.2	88	2	I40653	9K protein - Coxie
697	6	1.2	88	2	C82627	hypothetical prote
698	6	1.2	89	2	H72091	ribosomal prote
699	6	1.2	89	2	G86531	L28 ribosomal prot
700	6	1.2	89	2	T42967	hypothetical prote
701	6	1.2	89	2	G82606	hypothetical prote
702	6	1.2	90	2	B81102	conserved hypotet
703	6	1.2	90	2	G81135	hypothetical prote
704	6	1.2	90	2	E70004	hypothetical prote
705	6	1.2	90	2	AB3261	transglycosylase-a
706	6	1.2	90	2	AB2192	hypothetical prote
707	6	1.2	90	2	AH1955	hypothetical prote
708	6	1.2	91	1	S00060	phospholipid trans
709	6	1.2	91	2	AB1573	cardoxysome struct
710	6	1.2	91	2	AG1219	cardoxysome struct
711	6	1.2	91	2	E71938	hypothetical prote
712	6	1.2	91	2	AF2387	transglycosylase-a
713	6	1.2	92	2	DA3674	US5 protein - huma
714	6	1.2	92	2	E71868	hypothetical prote
715	6	1.2	92	2	I64109	mercury reductase
716	6	1.2	93	2	E97129	truncated flavodox
717	6	1.2	93	2	A95377	hypothetical prote
718	6	1.2	93	2	AF3485	hypothetical prote
719	6	1.2	94	2	AD1222	probable carboxys
720	6	1.2	94	2	AG1575	probable carboxys
721	6	1.2	94	2	A10759	probable propanedi
722	6	1.2	94	2	H82374	acetylactate synth
723	6	1.2	94	2	AG4863	hypothetical prote
724	6	1.2	95	2	G82866	phosphotransferase
725	6	1.2	95	2	AG1218	Salmonella typhimu
726	6	1.2	95	2	C75478	conserved hypotet
727	6	1.2	95	2	S72597	hypothetical prote
728	6	1.2	95	2	AG2439	hypothetical prote
729	6	1.2	96	2	D70254	hypothetical prote
730	6	1.2	96	2	C82970	hypothetical prote
731	6	1.2	96	2	SA0081	GPI-anchored lymph
732	6	1.2	97	2	T30458	hypothetical prote
733	6	1.2	97	2	T49767	hypothetical prote
734	6	1.2	97	2	D95328	hypothetical prote
735	6	1.2	98	2	S26940	Ig heavy chain V r
736	6	1.2	98	2	S38266	Ig heavy chain V r
737	6	1.2	98	2	P96643	cuticle protein lm
738	6	1.2	98	2	D75519	hypothetical prote
739	6	1.2	99	1	Q8BPM7	hypothetical prote
740	6	1.2	99	1	QOQYBW	host specificity p
741	6	1.2	99	2	S26800	hypothetical prote
742	6	1.2	99	2	S12418	Ig heavy chain V r
743	6	1.2	99	2	S12412	Ig heavy chain V r
744	6	1.2	99	2	B18115	Ig heavy chain V r
745	6	1.2	100	2	S78056	hypothetical prote
746	6	1.2	100	2	G70486	Ig heavy chain V r
747	6	1.2	100	2	A12465	integration host f
748	6	1.2	100	2	T06305	hypothetical prote
749	6	1.2	101	2	G71017	hypothetical prote
750	6	1.2	101	2	E87421	MADH-ubiquinone ox
751	6	1.2	101	2	S77990	hypothetical prote
752	6	1.2	102	2	S28361	hypothetical prote
753	6	1.2	102	2	S73799	probable lipoprote
754	6	1.2	102	2	T51524	hypothetical prote
755	6	1.2	103	2	E75334	probable dioxyena
756	6	1.2	103	2	F70178	hypothetical prote
757	6	1.2	104	2	S38268	cuticle protein lm
758	6	1.2	105	2	D89875	hypothetical prote
759	6	1.2	106	2	H32537	T-cell receptor be
760	6	1.2	106	2	G97584	hypothetical prote
761	6	1.2	106	2	AC2694	hypothetical prote
762	6	1.2	106	2	H97475	hypothetical prote
763	6	1.2	108	2	S01368	hypothetical prote
764	6	1.2	108	2	B72586	hypothetical prote
765	6	1.2	108	2	F82598	hypothetical prote
766	6	1.2	109	2	T01654	ribosomal protein
767	6	1.2	109	2	D72602	hypothetical prote
768	6	1.2	109	2	A72492	hypothetical prote
769	6	1.2	109	2	B86050	hypothetical prote
770	6	1.2	110	1	R6SSP2	acidic ribosomal p
771	6	1.2	110	1	R6BYP3	60s acidic ribosom
772	6	1.2	110	2	S44110	Ig heavy chain V-D
773	6	1.2	110	2	S76061	hypothetical prote
774	6	1.2	110	2	G87283	conserved hypotet
775	6	1.2	111	2	G91043	detoxt protein [imp
776	6	1.2	111	2	B85888	hypothetical prote
777	6	1.2	111	2	B85888	hypothetical prote
778	6	1.2	111	2	H55020	cchA protein [imp
779	6	1.2	111	2	T28220	hypothetical prote
780	6	1.2	111	2	B72722	hypothetical prote
781	6	1.2	111	2	A87525	EF hand domain pro
782	6	1.2	112	2	E90111	hypothetical prote
783	6	1.2	112	2	S54179	acidic ribosomal p
784	6	1.2	112	2	J01063	glycine-rich prote
785	6	1.2	113	1	R6FFP2	acidic ribosomal p
786	6	1.2	113	2	R6DOP1	acidic ribosomal p
787	6	1.2	113	2	P00273	polyprotein - hepa
788	6	1.2	113	2	G70754	hypothetical prote
789	6	1.2	114	2	S17387	T-cell receptor be
790	6	1.2	114	2	B34792	Ig heavy chain pre
791	6	1.2	114	2	T06885	ribosomal protein
792	6	1.2	114	2	AG1240	ribosome-binding f
793	6	1.2	114	2	DB4984	hypothetical prote
794	6	1.2	114	2	S23479	hypothetical prote
795	6	1.2	114	2	E72724	hypothetical prote
796	6	1.2	115	2	G69183	conserved hypotet
797	6	1.2	115	2	AH3305	nodulation protein
798	6	1.2	116	1	S66924	probable membrane
799	6	1.2	116	2	TPAF	insulin precursor
800	6	1.2	116	2	S21979	Ig heavy chain V-g
801	6	1.2	116	2	S17080	Ig heavy chain V-g
802	6	1.2	116	2	H71687	hypothetical prote
803	6	1.2	116	2	D95280	hypothetical prote
804	6	1.2	117	1	H95414	hypothetical prote
805	6	1.2	117	1	S60030	hemoglobin - Param
806	6	1.2	117	1	S60032	hemoglobin - Param
807	6	1.2	117	2	T02872	probable lipid tra
808	6	1.2	117	2	T09155	lipid transfer pro
809	6	1.2	117	2	E34792	Ig heavy chain pre
810	6	1.2	118	2	D90510	hypothetical prote
811	6	1.2	118	2	G34792	Ig heavy chain pre
812	6	1.2	118	2	D69837	hypothetical prote
813	6	1.2	119	2	PL0089	Ig heavy chain V r
814	6	1.2	119	2	AC1032	probable exported
815	6	1.2	119	2	T14932	hypothetical prote
816	6	1.2	119	2	AG1333	conserved hypotet
817	6	1.2	120	2	A34871	Ig kappa chain V r
818	6	1.2	120	2	PT0370	Ig mu chain precu
819	6	1.2	120	2	AD5598	hypothetical prote
820	6	1.2	121	2	T04093	phospholipid trans
821	6	1.2	122	2	AF1497	hypothetical prote
822	6	1.2	122	2	AG1139	hypothetical prote
823	6	1.2	122	2	S10213	hypothetical prote
824	6	1.2	122	2	T51036	hypothetical prote
825	6	1.2	122	2	F83167	hypothetical prote
826	6	1.2	122	2	C84320	hypothetical prote
827	6	1.2	123	2	E69980	hypothetical prote
828	6	1.2	123	2	F64083	ribosomal protein
829	6	1.2	123	2	H75430	conserved hypotet
830	6	1.2	123	2	A75273	hypothetical prote
831	6	1.2	123	2	T22970	hypothetical prote
832	6	1.2	124	1	H70633	hypothetical prote
					R5RFRA	ribosomal protein

541	7	1.4	847	2	T52436	614	7	1.4	3716	2	E70969	probable PPE prote
542	7	1.4	855	2	G86189	615	7	1.4	4196	2	T43274	dyein heavy chain
543	7	1.4	856	2	C71956	616	7	1.4	4613	2	T117409	polyketide synthas
544	7	1.4	856	2	H64552	617	7	1.4	4644	1	A38905	dyein heavy chain
545	7	1.4	871	2	A84177	618	7	1.4	5188	2	B85547	probable RTX fam1
546	7	1.4	886	2	A64112	619	7	1.4	5291	2	E90596	hypothetical prote
547	7	1.4	892	2	U00424	620	7	1.4	6260	2	T30328	polyketide synthas
548	7	1.4	895	2	T32780	621	7	1.4	8563	2	T30325	polyketide synthas
549	7	1.4	898	2	T37037	622	7	1.4	10223	2	T30325	polyketide synthas
550	7	1.4	915	2	AC3477	623	7	1.4	26926	15	S59489	titin cardiac mus
551	7	1.4	923	2	T38398	624	6	1.2	15	24	S59489	steroid monooxygen
552	7	1.4	928	2	T47846	625	6	1.2	29	24	PC2305	gaegurin 6 - Korea
553	7	1.4	941	2	T37626	626	6	1.2	33	2	A60683	malate dehydrogena
554	7	1.4	959	2	S32016	627	6	1.2	34	2	PO0408	RNA-directed RNA p
555	7	1.4	963	2	B83502	628	6	1.2	35	2	B42762	Ig heavy chain V r
556	7	1.4	966	2	G69189	629	6	1.2	39	2	S26939	C 3.4.25.1 proteas
557	7	1.4	969	2	AB3388	630	6	1.2	41	2	D55539	Ig heavy chain V r
558	7	1.4	974	1	A40213	631	6	1.2	42	2	S33485	hypothetical prote
559	7	1.4	987	2	I48373	632	6	1.2	44	2	S58438	hypothetical prote
560	7	1.4	992	1	GNMVR3	633	6	1.2	47	2	T40320	v1r-repressed prot
561	7	1.4	1001	2	G87385	634	6	1.2	47	2	PC4133	hypothetical 47 pr
562	7	1.4	1011	2	F70620	635	6	1.2	50	2	T14888	hypothetical prote
563	7	1.4	1012	2	T00958	636	6	1.2	51	2	T13196	hypothetical prote
564	7	1.4	1038	2	JC5497	637	6	1.2	55	2	E82522	hypothetical prote
565	7	1.4	1063	1	GNMW77	638	6	1.2	56	2	T44431	hypothetical prote
566	7	1.4	1066	2	G84746	639	6	1.2	57	2	T03654	glycerolaldehyde-3-p
567	7	1.4	1071	2	A84785	640	6	1.2	59	2	A86485	protein F2809.16 l
568	7	1.4	1079	2	F84946	641	6	1.2	59	2	H24802	citicic protein 64
569	7	1.4	1079	2	B70807	642	6	1.2	60	1	N1EY15	short neurotoxin 1
570	7	1.4	1091	2	S33596	643	6	1.2	60	1	N1EY1P	short neurotoxin 1
571	7	1.4	1104	2	A60999	644	6	1.2	60	1	N1AT1	short neurotoxin 1
572	7	1.4	1104	2	T49735	645	6	1.2	60	1	N1AT1F	short neurotoxin 1
573	7	1.4	1112	2	T02848	646	6	1.2	60	2	A27580	short neurotoxin 1
574	7	1.4	1122	2	B26427	647	6	1.2	61	2	S38917	hypothetical prote
575	7	1.4	1127	2	A25018	648	6	1.2	62	2	G25866	short neurotoxin b
576	7	1.4	1127	2	C26427	649	6	1.2	62	2	S14925	cerebellar polypep
577	7	1.4	1176	2	GNNVM7	650	6	1.2	62	2	A33915	hypothetical prote
578	7	1.4	1189	1	A40670	651	6	1.2	64	2	S36975	hypothetical prote
579	7	1.4	1201	2	A57369	652	6	1.2	64	2	H98033	hypothetical prote
580	7	1.4	1205	2	AH2486	653	6	1.2	65	2	T06457	outer envelope mem
581	7	1.4	1218	2	AC3580	654	6	1.2	65	2	AD0239	probable membrane
582	7	1.4	1218	2	A26427	655	6	1.2	66	2	B84278	hypothetical prote
583	7	1.4	1218	2	A26588	656	6	1.2	66	2	G82667	hypothetical prote
584	7	1.4	1221	2	A44978	657	6	1.2	66	2	A13117	hypothetical prote
585	7	1.4	1249	2	S68431	658	6	1.2	66	2	XSSOP5	excisionase - Stre
586	7	1.4	1280	2	A39117	659	6	1.2	67	1	E81042	hypothetical prote
587	7	1.4	1292	2	T09229	660	6	1.2	67	2	C72780	probable ribosomal
588	7	1.4	1343	2	AF0611	661	6	1.2	68	2	S7757	phosphotransferase
589	7	1.4	1361	2	E70806	662	6	1.2	69	2	B69355	hypothetical prote
590	7	1.4	1405	2	H87230	663	6	1.2	69	2	AC2563	hypothetical prote
591	7	1.4	1406	2	T17429	664	6	1.2	69	2	AP3566	hypothetical prote
592	7	1.4	1440	2	T44872	665	6	1.2	70	2	C53578	brevinin-1Bc precu
593	7	1.4	1473	2	T13855	666	6	1.2	71	2	LM2MA	H+-transporting tw
594	7	1.4	1486	2	F86785	667	74	1	74	1	LMVEM	H+-transporting tw
595	7	1.4	1515	2	T52081	668	6	1.2	74	1	LMWTM	H+-transporting tw
596	7	1.4	1533	2	A46321	669	6	1.2	74	2	S70027	H+-transporting tw
597	7	1.4	1538	2	H70646	670	6	1.2	74	2	S17767	H+-transporting tw
598	7	1.4	1557	2	T13160	671	6	1.2	74	2	S25960	H+-transporting tw
599	7	1.4	1640	2	D86798	672	6	1.2	74	2	A33667	H+-transporting tw
600	7	1.4	1686	2	A87692	673	6	1.2	74	2	AS9552	H+-transporting tw
601	7	1.4	1901	2	F70806	674	6	1.2	74	2	AG2533	H+-transporting tw
602	7	1.4	1938	1	A40997	675	6	1.2	75	2	I62129	hypothetical prote
603	7	1.4	1969	2	T08875	676	6	1.2	76	2	S58180	gene Mid1 protein
604	7	1.4	1978	2	T07081	677	6	1.2	76	2	AD3291	su11 protein - mou
605	7	1.4	1986	2	G86925	678	6	1.2	77	2	PC1017	hypothetical prote
606	7	1.4	2103	2	G86925	679	6	1.2	77	2	AC2805	somatotropin precu
607	7	1.4	2175	1	S03170	680	6	1.2	78	2	AB2548	hypothetical prote
608	7	1.4	2257	2	T09538	681	6	1.2	78	2	C95254	D-alanyl carrier p
609	7	1.4	2261	2	D86483	682	6	1.2	79	2	PN0641	hypothetical prote
610	7	1.4	2304	2	T07084	683	6	1.2	79	2	T42116	hypothetical prote
611	7	1.4	2359	2	E86483	684	6	1.2	79	2	A99519	D-alanyl carrier p
612	7	1.4	3002	2	A47221	685	6	1.2	79	2	AG3443	hypothetical prote
613	7	1.4	3436	2	S55659	686	6	1.2	79	2		

395	7	1.4	482	2	JC7583	468	7	1.4	591	2	T45713	hypothetical prote
396	7	1.4	483	2	AG1898	469	7	1.4	595	2	AP2632	conserved hypotnet
397	7	1.4	483	2	T21327	470	7	1.4	595	2	B97415	hypothetical prote
398	7	1.4	488	2	AB2984	471	7	1.4	595	2	B97531	branched-chain ami
399	7	1.4	489	2	C82282	472	7	1.4	595	2	AC2750	hypothetical prote
400	7	1.4	490	2	A96556	473	7	1.4	598	2	JN0573	polyadenylate-bind
401	7	1.4	496	2	H83592	474	7	1.4	598	2	T22610	hypothetical prote
402	7	1.4	502	2	A48679	475	7	1.4	599	2	F87682	TRKA family protei
403	7	1.4	503	2	C71550	476	7	1.4	603	2	AB7342	conserved hypotnet
404	7	1.4	503	2	T30966	477	7	1.4	603	2	B98107	hypothetical prote
405	7	1.4	504	2	T34995	478	7	1.4	606	2	T08180	p20 protein, micr
406	7	1.4	507	2	T35677	479	7	1.4	610	2	C96732	hypothetical prote
407	7	1.4	509	2	D84349	480	7	1.4	616	2	C83404	conserved hypotnet
408	7	1.4	511	2	G87609	481	7	1.4	617	2	T01227	glutathione-regula
409	7	1.4	520	2	S72324	482	7	1.4	623	2	T52293	MYC-related DNA bi
410	7	1.4	524	2	E82309	483	7	1.4	627	2	A41609	dnak-type molecula
411	7	1.4	526	2	E87548	484	7	1.4	627	2	T10290	hypothetical prote
412	7	1.4	526	2	AB6440	485	7	1.4	631	2	T05103	hypothetical prote
413	7	1.4	528	2	C96564	486	7	1.4	633	2	T27499	hypothetical prote
414	7	1.4	528	2	D46449	487	7	1.4	637	2	S48921	hypothetical prote
415	7	1.4	530	2	B87369	488	7	1.4	638	1	XXAV	dihydrolipoamide S
416	7	1.4	532	2	A72694	489	7	1.4	640	2	T08179	LRG5 protein - chl
417	7	1.4	533	2	S18539	490	7	1.4	643	2	T43052	fatty acid transpo
418	7	1.4	535	2	S07521	491	7	1.4	648	2	B40727	S-W checkpoint con
419	7	1.4	536	2	T37544	492	7	1.4	649	2	S48379	glycerol-3-phospha
420	7	1.4	541	2	T23088	493	7	1.4	654	2	T87602	sensory box histid
421	7	1.4	541	2	AB7595	494	7	1.4	654	2	T34613	NADH2 dehydrogenas
422	7	1.4	541	2	S73383	495	7	1.4	657	2	D71351	probable primosoma
423	7	1.4	542	2	S42089	496	7	1.4	659	2	T02838	probable membrane
424	7	1.4	542	2	AB2356	497	7	1.4	666	2	D87243	conserved integral
425	7	1.4	542	2	F97369	498	7	1.4	671	2	C96534	probable Poly-A Bi
426	7	1.4	542	2	AP2587	499	7	1.4	677	2	G87586	cation transportin
427	7	1.4	545	2	B89823	500	7	1.4	682	2	AG3351	transporter BME107
428	7	1.4	547	2	C98267	501	7	1.4	683	2	G82618	plus biogenesis p
429	7	1.4	547	2	AE3017	502	7	1.4	696	2	A90959	probable tail prot
430	7	1.4	548	1	STBYAL	503	7	1.4	696	2	C85807	hypothetical prote
431	7	1.4	549	1	D90477	504	7	1.4	696	2	A11029	hypothetical prote
432	7	1.4	553	2	JC7723	505	7	1.4	699	2	D70533	hypothetical prote
433	7	1.4	554	2	G75576	506	7	1.4	708	2	J01148	killer toxin KHS p
434	7	1.4	556	2	AD1770	507	7	1.4	733	2	D87112	primosomal protein
435	7	1.4	556	2	A11394	508	7	1.4	735	2	A59434	KIAA1501 protein
436	7	1.4	558	2	S52951	509	7	1.4	738	2	T00343	hypothetical prote
437	7	1.4	558	2	S52944	510	7	1.4	742	2	T35650	probable ATP-depen
438	7	1.4	559	2	S52941	511	7	1.4	749	2	B86606	primosomal protein
439	7	1.4	559	2	S52940	512	7	1.4	749	2	C72018	primosomal protein
440	7	1.4	559	2	S52937	513	7	1.4	750	2	T48804	hypothetical prote
441	7	1.4	559	2	S52950	514	7	1.4	752	2	T27276	hypothetical prote
442	7	1.4	559	2	S52935	515	7	1.4	752	2	AC3624	acridflavin resist
443	7	1.4	559	2	S52938	516	7	1.4	753	2	B71472	probable primosoma
444	7	1.4	559	2	S52956	517	7	1.4	753	2	D81734	primosomal protein
445	7	1.4	559	2	S52955	518	7	1.4	778	2	F70963	hypothetical glyci
446	7	1.4	559	2	S52954	519	7	1.4	778	2	D87965	protein V63D3A 6p
447	7	1.4	559	2	S52953	520	7	1.4	778	2	S56293	probable membrane
448	7	1.4	559	2	E84213	521	7	1.4	779	2	AB6862	primosomal protein
449	7	1.4	560	2	S52943	522	7	1.4	780	1	KIHUFM	6-phosphofructokin
450	7	1.4	560	2	H83206	523	7	1.4	780	1	KIRBF	6-phosphofructokin
451	7	1.4	564	2	F87305	524	7	1.4	780	2	S71429	phosphofructokin
452	7	1.4	566	2	T45162	525	7	1.4	781	2	A13014	methy1-accepting c
453	7	1.4	566	2	D70709	526	7	1.4	788	2	F98269	hypothetical prote
454	7	1.4	566	2	F71403	527	7	1.4	790	2	D83011	conserved hypotnet
455	7	1.4	570	1	C69985	528	7	1.4	795	2	B83608	hypothetical prote
456	7	1.4	570	2	A39597	529	7	1.4	797	2	AH1302	primosomal replica
457	7	1.4	570	2	S54134	530	7	1.4	797	2	AH1674	primosomal replica
458	7	1.4	571	2	S65060	531	7	1.4	798	2	C95202	primosomal protein
459	7	1.4	572	2	T00707	532	7	1.4	798	2	C98069	primosomal replica
460	7	1.4	573	2	C85433	533	7	1.4	802	2	G89893	primosomal p
461	7	1.4	574	2	S61234	534	7	1.4	803	2	B84931	DNA topoisomerase
462	7	1.4	575	2	B83586	535	7	1.4	804	2	E83963	primosomal replica
463	7	1.4	580	2	G83339	536	7	1.4	805	2	A69682	primosomal replica
464	7	1.4	583	2	T10701	537	7	1.4	813	2	AC3304	non-motile and pha
465	7	1.4	584	2	G70804	538	7	1.4	814	2	T14608	hypothetical prote
466	7	1.4	588	2	G95242	539	7	1.4	828	2	T03544	hypothetical prote
467	7	1.4	588	2	T22688	540	7	1.4	839	2	F75518	hypothetical prote

Tue Jan 28 10:21:47 2003

us-09-889-314-2.rpr

249	7	1.4	316	2	F85826	probable transcrip	322	7	1.4	398	1	OZ2QAS	circumsporozoite p
250	7	1.4	317	2	T00146	hypothetical prote	323	7	1.4	401	1	OZ2QAC	circumsporozoite p
251	7	1.4	317	2	AH0496	probable ornithine	324	7	1.4	401	2	AD1264	ammonium transport
252	7	1.4	320	2	T42062	thioredoxin-disulf	325	7	1.4	401	2	AF1626	ammonium transport
253	7	1.4	322	1	A53307	thioredoxin-disulf	326	7	1.4	404	2	H87106	probable integral
254	7	1.4	322	2	T36577	thioredoxin reduct	327	7	1.4	407	2	D84518	molYdenum colicoto
255	7	1.4	322	2	T40526	adp/atp translocas	328	7	1.4	407	2	D84169	hypothetical prote
256	7	1.4	323	2	S20099	transforming prote	329	7	1.4	410	2	S53933	probable membrane
257	7	1.4	324	2	F90311	hypothetical prote	330	7	1.4	412	2	F93394	probable MFS trans
258	7	1.4	324	2	AH2844	thioredoxin reduct	331	7	1.4	413	2	B83180	hypothetical prote
259	7	1.4	324	2	T44374	thioredoxin reduct	332	7	1.4	414	2	S52618	hypothetical prote
260	7	1.4	328	2	AF2081	conserved hypothet	333	7	1.4	415	2	T36906	probable integral
261	7	1.4	330	2	B87502	flagellar motor sw	334	7	1.4	416	1	A41267	transcription fact
262	7	1.4	331	1	A30929	flagellar motor sw	335	7	1.4	416	1	E83590	cytosine permease
263	7	1.4	331	2	H64957	flagellar motor sw	336	7	1.4	417	2	AD0907	cytosine permease
264	7	1.4	331	2	F90963	hypothetical prote	337	7	1.4	417	2	E90677	cytosine permease
265	7	1.4	331	2	F85811	flagellar motor sw	338	7	1.4	419	2	A83528	cytosine permease
266	7	1.4	331	2	AC0752	hypothetical glycol	339	7	1.4	419	2	S22651	cytosine permease/
267	7	1.4	331	2	F70820	high-affinity bran	340	7	1.4	420	2	S06579	finger protein (cl
268	7	1.4	331	2	AC3552	hypothetical prote	341	7	1.4	420	2	T50585	probable membrane
269	7	1.4	332	1	S44027	thioredoxin-disulf	342	7	1.4	420	2	E86162	probable membrane
270	7	1.4	333	1	D70716	thioredoxin-disulf	343	7	1.4	422	2	D84403	hypothetical prote
271	7	1.4	333	2	T41743	homeotic protein, e	344	7	1.4	422	2	H82418	hypothetical prote
272	7	1.4	333	2	S53596	chlorophyll a/b-bi	345	7	1.4	423	2	B70931	ATP-dependent RNA
273	7	1.4	335	2	S53596	chlorophyll a/b-bi	346	7	1.4	425	2	A31026	probable PPE prote
274	7	1.4	337	2	T06196	hypothetical prote	347	7	1.4	426	1	Z4BPD	probable membrane
275	7	1.4	339	2	A83128	acetoIn catabolism	348	7	1.4	426	1	Z4BPF1	gene IV protein -
276	7	1.4	339	2	T43925	ABC transporter in	349	7	1.4	426	1	T35676	probable ABC-type
277	7	1.4	344	2	A70771	probable hydrolase	350	7	1.4	430	2	H75622	hypothetical prote
278	7	1.4	345	2	A64493	acyl carrier prote	351	7	1.4	432	2	B84566	hypothetical prote
279	7	1.4	346	1	MUHUGP	N4-(beta-N'-acetyl	352	7	1.4	433	2	U01459	conserved hypothet
280	7	1.4	346	2	AD3210	ABC transporter, s	353	7	1.4	436	2	E70387	conserved hypothet
281	7	1.4	347	2	S44846	KO6H7, 6 protein -	354	7	1.4	438	2	JC7758	OYCAL protein - mo
282	7	1.4	347	2	D97588	bifunctional carbo	355	7	1.4	439	2	B82114	hypothetical prote
283	7	1.4	351	2	D64185	UDP-N'-acetylglucos	356	7	1.4	440	2	G87354	hypothetical prote
284	7	1.4	353	2	B70358	hydrogenase (EC 1.	357	7	1.4	441	2	A73661	hypothetical prote
285	7	1.4	358	2	H97613	hypothetical prote	358	7	1.4	445	2	T36069	hypothetical prote
286	7	1.4	358	2	AE2836	lytic murein trans	359	7	1.4	446	2	AE0088	WSC4 homolog [limp
287	7	1.4	359	2	I49341	MIP-1 alpha recept	360	7	1.4	446	2	T45525	H+-transporting tw
288	7	1.4	359	2	T26813	hypothetical prote	361	7	1.4	450	2	C87463	hypothetical prote
289	7	1.4	361	2	H72587	hypothetical prote	362	7	1.4	451	2	G82096	sodium-dependent t
290	7	1.4	362	2	S45094	cinamyl-alcohol d	363	7	1.4	451	2	S76671	hypothetical prote
291	7	1.4	364	2	C48376	orf 5' to pbc -	364	7	1.4	452	2	F63906	conserved hypothet
292	7	1.4	365	2	T12040	cysteine proteinas	365	7	1.4	452	2	T16062	hypothetical prote
293	7	1.4	365	2	F96986	sugar ABC transpor	366	7	1.4	454	2	T26296	hypothetical prote
294	7	1.4	366	2	T16119	hypothetical prote	367	7	1.4	454	2	H70820	hypothetical glycol
295	7	1.4	366	2	T31820	hypothetical prote	368	7	1.4	457	2	B64818	glucuronide permea
296	7	1.4	366	2	A97451	GTP cyclohydrolase	369	7	1.4	457	2	C90919	glucuronide permea
297	7	1.4	366	2	AC2669	hypothetical prote	370	7	1.4	457	2	H85767	thioredoxin-disulf
298	7	1.4	367	2	G71178	homeobox protein -	371	7	1.4	458	1	S77662	dihydrodipionamide a
299	7	1.4	368	2	T42295	hypothetical prote	372	7	1.4	458	2	A30775	conserved hypothet
300	7	1.4	369	2	S60187	peridin-in-chloroph	373	7	1.4	462	2	H82160	probable narv prot
301	7	1.4	370	2	AC3440	GTP cyclohydrolase	374	7	1.4	463	2	F70627	hypothetical prote
302	7	1.4	371	2	B82933	hypothetical prote	375	7	1.4	463	2	E70976	hypothetical prote
303	7	1.4	375	2	B82933	nitrite reductase	376	7	1.4	467	2	G72602	hypothetical prote
304	7	1.4	376	2	A53384	nitrite reductase	377	7	1.4	469	2	S30830	hypothetical prote
305	7	1.4	377	2	A53384	polyisaltic acid ca	378	7	1.4	470	1	AF0771	hypothetical prote
306	7	1.4	377	2	S60758	stiaa protein - Nel	379	7	1.4	470	2	AH3594	flavinate transpor
307	7	1.4	378	2	UC4648	nitrite reductase	380	7	1.4	474	2	S52444	flavinate protein
308	7	1.4	379	2	A48936	nitrite reductase	381	7	1.4	474	2	S52444	flavinate protein
309	7	1.4	379	2	T35653	hypothetical prote	382	7	1.4	474	2	S16250	flavinate protein
310	7	1.4	382	2	AE2754	probable acyl-CoA	383	7	1.4	474	2	F85430	phytoene dehydroge
311	7	1.4	382	2	C97535	two component sens	384	7	1.4	475	2	T03170	hypothetical prote
312	7	1.4	382	2	C97535	nitrogen regulatio	385	7	1.4	475	2	G64720	hypothetical prote
313	7	1.4	383	2	A84552	probable thioredox	386	7	1.4	476	2	G90629	hypothetical prote
314	7	1.4	384	2	S25771	gas1 protein - mou	387	7	1.4	476	2	G85480	hypothetical prote
315	7	1.4	388	2	F75162	tryptophan synthas	388	7	1.4	476	2	D75469	probable inner mem
316	7	1.4	388	2	AF1907	tryptophan synthas	389	7	1.4	477	2	D75469	probable inner mem
317	7	1.4	389	2	T43927	tryptophan synthas	390	7	1.4	477	2	T12683	embryogenesis prot
318	7	1.4	391	2	A40594	major flagellin pr	391	7	1.4	478	2	F87474	major facilitator
319	7	1.4	392	2	C72427	pyruvate synthase	392	7	1.4	479	2	AB2035	pyruvate desaturas
320	7	1.4	392	2	C72427	immobilization sur	393	7	1.4	481	2	T00863	hypothetical prote
321	7	1.4	396	2	T45088	pyruvate synthase	394	7	1.4	481	2	T00863	hypothetical prote

103	7	1.4	85	2	C87715	hypothetical prote
104	7	1.4	92	1	B87456	ribosomal protein
105	7	1.4	93	1	A37086	small acid-soluble
106	7	1.4	98	2	H95391	probable thiooxid
107	7	1.4	100	2	G97672	hypothetical prote
108	7	1.4	100	2	AE2897	conserved hypotet
109	7	1.4	106	1	R8BY28	acidic ribosomal p
110	7	1.4	107	2	S03703	12K protein - pea
111	7	1.4	110	1	R5BYA1	acidic ribosomal p
112	7	1.4	110	2	T37490	ribosomal protein
113	7	1.4	110	2	T01565	acidic ribosomal p
114	7	1.4	112	2	S35440	acidic ribosomal p
115	7	1.4	113	2	S43109	acidic ribosomal p
116	7	1.4	113	2	P00274	polypeptide - hepa
117	7	1.4	118	2	D72524	hypothetical prote
118	7	1.4	118	2	F87447	hypothetical prote
119	7	1.4	120	2	AD1106	ribosomal protein
120	7	1.4	122	1	R5HG12	ribosomal protein
121	7	1.4	122	1	T30865	stbC protein - Sal
122	7	1.4	123	1	R5BS9	ribosomal protein
123	7	1.4	123	2	G97286	ribosomal protein
124	7	1.4	123	2	D64003	probable sodium-tr
125	7	1.4	124	2	I40348	ribosomal protein
126	7	1.4	124	2	I40350	ribosomal protein
127	7	1.4	124	2	AE3345	LSU ribosomal prot
128	7	1.4	126	2	G70685	hypothetical prote
129	7	1.4	127	2	H87310	ribosomal protein
130	7	1.4	130	2	D84353	hypothetical prote
131	7	1.4	130	2	E65042	hypothetical prote
132	7	1.4	131	2	S75262	sensory transducti
133	7	1.4	132	2	P00394	genome polypeptei
134	7	1.4	132	2	P00396	genome polypeptei
135	7	1.4	133	2	T31324	hypothetical prote
136	7	1.4	138	2	C87389	hypothetical prote
137	7	1.4	140	2	D97722	50S ribosomal prot
138	7	1.4	142	2	B32998	chorion protein S1
139	7	1.4	146	2	S15449	hemoglobin beta ch
140	7	1.4	148	2	S50653	hypothetical prote
141	7	1.4	149	2	H75462	hypothetical prote
142	7	1.4	152	2	T49126	hypothetical prote
143	7	1.4	154	2	S04440	fibrial protein -
144	7	1.4	154	2	B31105	fibrial protein p
145	7	1.4	155	2	D87414	hypothetical prote
146	7	1.4	157	2	T02034	early light-induce
147	7	1.4	163	1	D84674	hypothetical prote
148	7	1.4	165	1	C009	plastoeyanin precu
149	7	1.4	167	2	T04903	iron-sulfur cofact
150	7	1.4	168	2	T02036	early light-induce
151	7	1.4	168	2	G72679	hypothetical prote
152	7	1.4	171	2	S76927	hypothetical prote
153	7	1.4	171	2	B87628	conserved hypotet
154	7	1.4	173	2	AD0158	probable Mut faml
155	7	1.4	176	2	F71540	hypothetical prote
156	7	1.4	178	2	E77069	hypothetical prote
157	7	1.4	178	2	E82742	hypothetical prote
158	7	1.4	179	2	D85602	probable complemen
159	7	1.4	180	2	G82858	hypothetical prote
160	7	1.4	184	2	S12095	embryonic abundant
161	7	1.4	185	2	F84182	hypothetical prote
162	7	1.4	188	2	G96652	protein F23N19.15
163	7	1.4	190	2	C71421	hypothetical prote
164	7	1.4	192	2	H91103	probable transport
165	7	1.4	192	2	C85949	hypothetical prote
166	7	1.4	192	2	AC0875	probable membrane
167	7	1.4	194	2	S24297	DNA-3-methyladeni
168	7	1.4	195	2	E87296	hypothetical prote
169	7	1.4	200	2	H83297	LEA protein - rice
170	7	1.4	200	2	T04147	hypothetical prote
171	7	1.4	205	2	F87623	conserved hypotet
172	7	1.4	206	2	AC0692	conserved hypotet
173	7	1.4	207	2	I38602	zinc finger protei
174	7	1.4	215	2	E86656	transcription regu
175	7	1.4	215	2	D84382	hypothetical prote
176	7	1.4	216	2	F69166	hypothetical prote
177	7	1.4	220	2	S30977	gene 32 protein -
178	7	1.4	225	2	A87624	hypothetical prote
179	7	1.4	225	2	T50707	uracase accessory p
180	7	1.4	225	2	C70500	probable transcrip
181	7	1.4	227	1	SKXAD8	hexon-associated p
182	7	1.4	227	1	SKXAD5	hexon-associated p
183	7	1.4	229	1	C43330	gene 7 protein - p
184	7	1.4	230	2	G89892	orotidine-5-phosph
185	7	1.4	230	2	S72714	Lepb1170_F2_64 pro
186	7	1.4	231	2	T21119	hypothetical prote
187	7	1.4	233	2	S33946	hexon-associated p
188	7	1.4	235	2	T19328	hypothetical prote
189	7	1.4	237	2	H75374	beta-phosphoglucom
190	7	1.4	243	2	S07997	lipoprotein trar p
191	7	1.4	243	2	A38901	lipoprotein trar p
192	7	1.4	243	2	G64843	probable peroxidase
193	7	1.4	244	2	S01757	lipoprotein trar p
194	7	1.4	244	2	F70834	hypothetical prote
195	7	1.4	244	2	AB3609	multidrug resistanc
196	7	1.4	245	2	AE2195	hypothetical prote
197	7	1.4	248	2	H80792	hypothetical prote
198	7	1.4	248	2	G75214	methyltransferase
199	7	1.4	249	2	G81218	bioH protein, prob
200	7	1.4	251	2	AB1751	triose phosphate 1
201	7	1.4	251	2	A11381	triose phosphate 1
202	7	1.4	251	2	S20455	pgqc protein - Kle
203	7	1.4	259	2	AF3174	3'-oxoacyl-(acyl)-ca
204	7	1.4	260	2	T35293	probable cobalamin
205	7	1.4	262	2	S00275	tail fiber protein
206	7	1.4	266	2	P00393	genome polypeptei
207	7	1.4	267	2	B87578	hypothetical prote
208	7	1.4	267	2	E72690	probable spermidin
209	7	1.4	269	2	T24358	hypothetical prote
210	7	1.4	269	2	E97296	pUR operon repress
211	7	1.4	271	1	A34950	dihydropterate redu
212	7	1.4	272	2	G96930	nitrogenase iron p
213	7	1.4	276	2	E72643	probable phospho-2
214	7	1.4	276	2	G90786	probable cytochrom
215	7	1.4	277	2	T44589	tyrosin biosynthes
216	7	1.4	278	2	T08647	hypothetical prote
217	7	1.4	279	2	T48106	hypothetical prote
218	7	1.4	279	2	G85646	hypothetical prote
219	7	1.4	281	2	AB1585	B. subtilis yxhD p
220	7	1.4	282	2	AC0226	probable membrane
221	7	1.4	283	1	A56649	dihydropterate sy
222	7	1.4	283	2	A75274	hypothetical prote
223	7	1.4	284	2	G72662	hypothetical prote
224	7	1.4	287	2	A54601	RNA-binding protei
225	7	1.4	289	2	T49154	DNA-binding WRKY-1
226	7	1.4	291	2	T34043	hypothetical prote
227	7	1.4	298	2	C96022	probable 3-hydroxy
228	7	1.4	299	2	T35765	hypothetical prote
229	7	1.4	300	2	S24057	ferritin 2 precurs
230	7	1.4	300	2	B75619	probable cobalamin
231	7	1.4	300	2	E84310	hypothetical prote
232	7	1.4	302	1	S44026	thioredoxin-disulf
233	7	1.4	304	2	AH0763	probable transcrip
234	7	1.4	306	2	T42011	ADP,ATP carrier pr
235	7	1.4	306	2	S08607	chorion protein s3
236	7	1.4	307	2	AB2029	ribokinase [import
237	7	1.4	309	2	A12809	hypothetical prote
238	7	1.4	312	2	C81795	probable hydrolase
239	7	1.4	312	2	F70044	hypothetical prote
240	7	1.4	313	2	A75468	4-hydroxybenzoate
241	7	1.4	313	2	E91006	penicillin-binding
242	7	1.4	313	2	F85850	penicillin-binding
243	7	1.4	313	2	E64981	penicillin-binding
244	7	1.4	314	2	C98193	sugar transport sy
245	7	1.4	314	2	AC3093	hypothetical prote
246	7	1.4	316	1	F64966	probable transcrip
247	7	1.4	316	2	B97622	thioredoxin reduct
248	7	1.4	316	2	A99981	probable transcrip

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 16:34:53 ; Search time 23 Seconds
(without alignments)
2073.160 Million cell updates/sec

Title: US-09-889-314-2
Perfect score: 496

Sequence: 1 DTWMSISSSGPDNCKNIMS.....LKAYAAISGATAGAKTKTNF 496

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 283224 seqs, 96134422 residues

Word size : 0
Total number of hits satisfying chosen parameters: 283224

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: listing first 1000 summaries

Database : PIR73:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	493	99.4	493	2 F72031	CT578 hypothetical
2	487	98.2	488	2 A81507	conserved hypothet
3	487	98.2	488	2 G86591	CT578 hypothetical
4	11	2.2	2038	2 A43742	female sterile hom
5	10	2.0	128	2 T30714	hypothetical prote
6	9	1.8	111	2 S43115	acidic ribosomal p
7	9	1.8	153	2 T49895	oleosin-like prote
8	9	1.8	169	2 T38724	mitochondrial benz
9	9	1.8	169	2 S14257	benzodiazepine rec
10	9	1.8	169	2 I38105	peripheral benzodi
11	9	1.8	169	2 JE0149	peripheral benzodi
12	9	1.8	169	2 T34870	secretion protein,
13	9	1.8	437	2 AE2972	rhizobium secret
14	9	1.8	474	2 P98310	hypothetical prote
15	9	1.8	487	2 D71497	conserved hypothet
16	9	1.8	491	2 F81655	elastin precursor,
17	9	1.8	747	1 EABO	elastin precursor,
18	9	1.8	770	2 S59623	tropoelastin - she
19	9	1.8	860	1 EAMS	elastin precursor
20	9	1.8	967	2 I40889	potassium channel
21	9	1.8	1001	2 T13807	probable polyketid
22	9	1.8	1017	2 B70985	dye protein - frui
23	9	1.8	1019	2 T00117	polyketide synthas
24	9	1.8	1616	2 G70668	hypothetical prote
25	9	1.8	2484	2 T26216	hypothetical prote
26	9	1.8	2607	2 T26215	fibroin - Chinese
27	9	1.8	2639	2 T31328	hypothetical prote
28	8	1.6	81	2 T35731	ribosomal protein
29	8	1.6	106	2 T52147	

30	8	1.6	110	1 R6BY22	60s acidic ribosom
31	8	1.6	110	1 R6BY24	60s acidic ribosom
32	8	1.6	111	1 R8S512	ribosomal protein
33	8	1.6	111	1 S41866	acidic ribosomal p
34	8	1.6	112	1 R6UT2B	acidic ribosomal p
35	8	1.6	132	2 G70354	hypothetical prote
36	8	1.6	143	2 A83149	hypothetical prote
37	8	1.6	145	2 H75262	hypothetical prote
38	8	1.6	161	1 S43893	hypothetical prote
39	8	1.6	161	1 S32970	H+-exporting ATPas
40	8	1.6	165	1 A40814	H+-exporting ATPas
41	8	1.6	178	2 G87307	hypothetical prote
42	8	1.6	178	2 E88637	hypothetical prote
43	8	1.6	179	2 A85217	hypothetical prote
44	8	1.6	209	2 E70536	hypothetical prote
45	8	1.6	214	2 G97631	conserved hypothet
46	8	1.6	214	2 AB2855	hypothetical prote
47	8	1.6	222	2 T50958	probable membrane
48	8	1.6	254	2 T35388	hypothetical prote
49	8	1.6	262	2 H70953	hypothetical prote
50	8	1.6	272	2 D36802	hypothetical prote
51	8	1.6	277	2 T04441	IR6 protein - equi
52	8	1.6	281	2 E72759	hypothetical prote
53	8	1.6	303	2 H69551	hypothetical prote
54	8	1.6	305	2 JC5844	acyl carrier prote
55	8	1.6	311	2 A42982	chitinase (EC 3.2.
56	8	1.6	313	2 A41524	rfas protein - Esc
57	8	1.6	338	2 S42640	transcription fact
58	8	1.6	354	2 T46740	photosystem II 30
59	8	1.6	361	2 T39529	microfilament shea
60	8	1.6	362	2 T41842	hyal protein - Al
61	8	1.6	362	2 C70858	ACMNPV orf101 - Bo
62	8	1.6	379	2 H87458	hypothetical prote
63	8	1.6	401	2 C83109	conserved hypothet
64	8	1.6	407	2 H71857	probable transport
65	8	1.6	407	2 F64658	pyruvate synthase
66	8	1.6	416	2 T37023	pyruvate synthase
67	8	1.6	434	2 E84879	probable oxidoredu
68	8	1.6	450	2 B97297	probable heme A fa
69	8	1.6	469	2 T01579	hydrogenase chain
70	8	1.6	471	2 E96573	heme A farnesyltra
71	8	1.6	472	2 S74886	protein F12M16.23
72	8	1.6	479	2 AG2103	phytoene dehydroge
73	8	1.6	491	2 C84362	zeta-carotene desa
74	8	1.6	504	2 E56542	Na+/H+ antiporter
75	8	1.6	512	2 AC1323	calmodulin-binding
76	8	1.6	512	2 AD1694	2-isopropylmalate
77	8	1.6	512	2 H89152	protein G24B5.2 li
78	8	1.6	587	2 C86174	myosin-crossreacti
79	8	1.6	594	2 G70545	hypothetical glyci
80	8	1.6	625	2 T03922	hypothetical prote
81	8	1.6	631	2 JC5803	calcium/calmodulin
82	8	1.6	632	2 JC7155	ring finger protei
83	8	1.6	650	2 S75072	brain finger protei
84	8	1.6	687	2 B57713	probable phytoene
85	8	1.6	718	2 A36068	chloride channel k
86	8	1.6	733	2 S44876	major ampullare fi
87	8	1.6	798	2 T21369	hypothetical glyci
88	8	1.6	853	2 A70896	hypothetical prote
89	8	1.6	892	2 AB7447	sensor histidine k
90	8	1.6	892	2 T09193	ataxin 7 - human
91	8	1.6	1251	2 C82721	conserved hypothet
92	8	1.6	1337	2 B64993	hypothetical prote
93	8	1.6	1337	2 B64993	probable calcium-b
94	8	1.6	1597	1 S45781	gene small optic l
95	8	1.6	1698	2 T13800	coracle gene prote
96	8	1.6	2129	2 T14182	icbc protein - Myc
97	8	1.6	2569	2 T14164	peptidase synthetase
98	8	1.6	75	2 T29665	hypothetical prote
99	8	1.4	76	2 H70966	hypothetical prote
100	7	1.4	76	2 G64663	ribosomal protein
101	7	1.4	80	2 E71430	hypothetical prote
102	7	1.4	80	2 E71430	

THIS PAGE BLANK (USPTO)

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID-7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-OREGON R;
 RA Nakagoshi H., Hoshi M., Nabeshima Y., Matsuzaki F.;
 RT "A novel homeobox gene mediates the DPP signal to establish functional
 RT specificity within target cells.";
 RL Genes Dev. 0:0-0(1998).
 DR EMBL: AB010299; BAA32660.1; -.
 DR HSSP: P10037; 1A07.
 DR FlyBase: FBgn0020307; dve.
 DR InterPro: IPR001356; Homeobox.
 DR Pfam: PF00046; homeobox; 3.
 DR SMART: SM00389; HOX; 2.
 DR PROSITE: PS50071; HOMEBOX_2; 2.
 SQ SEQUENCE 1019 AA; 11135 MW; 5D22A104A0D0F0CC CRC64;

Query Match 1.88; Score 9; DB 5; Length 1019;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 AAAGAAG 302 |||||
 DB 477 AAAGAAG 485

Search completed: January 27, 2003, 16:36:18
 Job time : 56 secs

OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 RX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Smith D.R.;
 RL Submitted (JAN-1994) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Robison K.;
 RL Submitted (SEP-1994) to the EMBL/GenBank/DBJ databases.
 DR EMBL: U00024; AAA50928.1; -.
 DR HSSP; P25715; IMMA.
 DR InterPro: IPR001227; Ac.transferase.
 DR InterPro: IPR001680; WD40.
 DR Pfam; PF00698; Acyl_transf_1.
 DR PROSITE; PS00678; WD_REPEATS_1; UNKNOWN_1.
 FT NON_TER 950 950
 SQ SEQUENCE 950 AA; 97188 MW; 305507720FB50DD8 CRC64;

Query Match 1.8%; Score 9; DB 2; Length 950;
 Best Local Similarity 100.0%; Pred. No. 22;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 AGLAGIAG 289
 DB 27 AGLAGIAG 35

RESULT 43
 0960M0 PRELIMINARY; PRT; 1013 AA.
 AC 0960M0;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE LD45559P.
 GN CG1869.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 CC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 CC Ephydroidea; Drosophilidae; Drosophila.
 RX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Aghayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacleb J., Paragas V., Park S., Phouenavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Geulker S.;
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY051988; AAK83412.1; -.
 DR FlyBase; FBgn0035398; CG1869.
 DR InterPro: IPR001579; Chitinase_18/2.
 DR InterPro: IPR002557; Chitin_bind_Pera.
 DR InterPro: IPR001223; Glyco_hydro_18.
 DR Pfam; PF01607; CBM_14; 1.
 DR Pfam; PF00704; Glyco_hydro_18; 2.
 DR PRODOM; PD000471; Glyco_hydro_18; 2.
 DR PROSITE; PS01095; CHITINASE_18; UNKNOWN_2.
 SQ SEQUENCE 1013 AA; 113285 MW; 266981AD0685EC02 CRC64;

Query Match 1.8%; Score 9; DB 5; Length 1013;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGAA 301
 DB 88 GAAAGGAA 96

RESULT 44
 006586

ID 006586 PRELIMINARY; PRT; 1017 AA.
 AC 006586;
 DT 01-JUL-1997 (TREMBLrel. 04, Created)
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Hypothetical 106.4 kDa protein (polyketide synthase).
 GN PKS9 OR RV1664 OR MTCY275.03 OR MT1704.
 OS Mycobacterium tuberculosis.
 OC Actinomycetales; Actinobacteria; Actinobacteridae;
 CC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
 RX NCBI_TaxID=1773;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=H37RV;
 RX MEDLINE=98295987; PubMed=9634230;
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
 RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
 RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;
 RT "Deciphering the biology of Mycobacterium tuberculosis from the
 RT complete genome sequence.";
 RL Nature 393:537-544(1998).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CDC 1551 / OSHKOSH;
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
 RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
 RA Bishai W., Utterback T., Weidman J., Knout H., Gill J., Mikula A.,
 RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
 RT laboratory strains.";
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; Z95617; CAB09100.1; -.
 DR EMBL; AE007033; AAK45971.1; -.
 DR HSSP; P25715; IMMA.
 DR TIGR; MT1704; -.
 DR TUBerculist; RV1664; -.
 DR InterPro: IPR001227; Ac.transferase.
 DR InterPro: IPR000794; ketoacyl-synt.
 DR InterPro: IPR003880; Panine. attach.
 DR Pfam; PF00698; Acyl_transf_1.
 DR Pfam; PF00109; ketoacyl-synt; 1.
 DR Pfam; PF02801; ketoacyl-synt_C; 1.
 DR Pfam; PF00550; pp-binding; 1.
 DR PROSITE; PS00075; ACP_DOMAIN; 1.
 DR PROSITE; PS00012; PHOSPHOPANTHETINE; UNKNOWN_1.
 KW Hypothetical protein; Phosphopantetheine; Complete proteome.
 SQ SEQUENCE 1017 AA; 106402 MW; B93A22D8818AE0EC CRC64;

Query Match 1.8%; Score 9; DB 16; Length 1017;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 281 AGLAGIAG 289
 DB 478 AGLAGIAG 486

RESULT 45
 076143 PRELIMINARY; PRT; 1019 AA.
 AC 076143;
 DT 01-NOV-1998 (TREMBLrel. 08, Created)
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE DVE.
 GN DVE OR CG5799.
 OS Drosophila melanogaster (Fruit fly).

```

Query Match          1.8%; Score 9; DB 5; Length 696;
Best Local Similarity 100.0%; Pred. No. 17;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGCAA 301
   |||||||
DB 88 GAAAGCAA 96

RESULT 39
O9LID0 PRELIMINARY; PRT; 746 AA.
AC O9LID0;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE Similarity to amine oxidase.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20363099; PubMed=10907853;
RA Kaneko T., Kato T., Sato S., Nakamura Y., Asamizu E., Tabata S.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=COLUMBIA;
RX MEDLINE=20363099; PubMed=10907853;
RA Nakamura Y.;
RT "Structural analysis of Arabidopsis thaliana chromosome 3. II.
RT Sequence features of the regions of 4,251,695 bp covered by ninety pl,
RT "nc and BAC clones.";
RL DNA Res. 7:217-221(2000).
DR EMBL; AP001307; BAB01917.1; -
DR InterPro: IPR002937; Amino-oxidase.
DR InterPro: IPR000205; NAD_binding.
DR Pfam: PF01593; Amino_oxidase; 1.
SQ SEQUENCE 746 AA; 82350 MW; E0FDCC0119C4EF00 CRC64;

Query Match          1.8%; Score 9; DB 10; Length 746;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 280 GAGLAGTAA 288
   |||||||
DB 166 GAGLAGTAA 174

RESULT 40
O8W0N1 PRELIMINARY; PRT; 748 AA.
AC O8W0N1;
DT 01-MAR-2002 (TREMBLrel. 20, Created)
DT 01-MAR-2002 (TREMBLrel. 20, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical 82.1 kDa protein.
CN P0681B1.1 OR B1085F01.16.
OS Oryza sativa (Rice), and
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
OX NCBI_TaxID=4530, 39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0681B1.";
```

```

RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
RT clone:B1085F01.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003022; BAB/8635.1; -
DR EMBL; AP003330; BAB89980.1; -
DR InterPro: IPR000719; Euk_Pkinase.
DR InterPro: IPR002290; Ser_thr_Pkinase.
DR InterPro: IPR001245; Tyr_Pkinase.
DR Pfam; PF00069; Pkinase; 1.
DR ProDom; PD000001; Euk_Pkinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TYRKC; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00109; PROTEIN_KINASE_TYR; UNKNOWN_1.
KW Hypothetical protein.
SQ SEQUENCE 748 AA; 82079 MW; E18542CC048E36F7 CRC64;

Query Match          1.8%; Score 9; DB 10; Length 748;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 298 GGAAGAAA 306
   |||||||
DB 23 GGAAGAAA 31

RESULT 41
O9ESZ9 PRELIMINARY; PRT; 810 AA.
AC O9ESZ9;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE ELN (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129/SV;
RA Green E.D.;
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF289665; AAF99336.1; -
DR InterPro: IPR001899; Gram_pos_anchor.
DR PROSITE; PS00343; GRAM_POS_ANCHORING; UNKNOWN_1.
FT NON_TER 1
FT 1
SQ SEQUENCE 810 AA; 67204 MW; 1ED61ED5AD4F40A2 CRC64;

Query Match          1.8%; Score 9; DB 11; Length 810;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 298 GGAAGAAA 306
   |||||||
DB 63 GGAAGAAA 71

RESULT 42
O50470 PRELIMINARY; PRT; 950 AA.
AC O50470;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE PK8002C (Fragment).
OC Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
```

Query Match 1.8%; Score 9; DB 4; Length 555;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 280 GAGLAGLAA 288
 |||||
 DB 31 GAGLAGLAA 39

RESULT 36

OY 099K82 PRELIMINARY; PRT; 555 AA.
 AC 099K82;
 DT 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Similar to hypothetical protein.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.

RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC004831; AA04831.1; -
 DR InterPro: IPR000759; Adrnck_reductase.
 DR InterPro: IPR002937; Amino_oxidase.
 DR InterPro: IPR000205; NAD_binding.
 DR Pfam: PF01593; Amino_oxidase; 1.
 DR PRINTS: PR00419; ADXRDTASE.
 KM Hypothetical protein.
 SQ SEQUENCE 555 AA; 61852 MW; A297E9DBD094EA74 CRC64;

Query Match 1.8%; Score 9; DB 11; Length 555;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 280 GAGLAGLAA 288
 |||||
 DB 31 GAGLAGLAA 39

RESULT 37

OY 091TN6 PRELIMINARY; PRT; 575 AA.

AC 091TN6;
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE T52.
 OS Tupaiia herpesvirus.
 OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
 OC Betaherpesvirinae.
 OX NCBI_TaxID=10397;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=2;
 RA MEDLINE=21211637; PubMed=11312357;
 RA Bahr U.; Darai G.;
 RT "Analysis and Characterization of the Complete Genome of Tupaiia (Tree
 RT Shrew) Herpesvirus."
 RL J. Virol. 75:4854-4870(2001).
 RN [2]
 RP SEQUENCE FROM N.A.

RC STRAIN=2;
 RA Darai G.; Bahr U.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF281817; AAK57101.1; -
 DR InterPro: IPR002597; Herpes_env.
 DR Pfam: PF01673; Herpes_env; 1.
 SQ SEQUENCE 575 AA; 63743 MW; BE62EEF35166825D CRC64;

Query Match 1.8%; Score 9; DB 12; Length 575;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 299 GAGGAAAT 307
 |||||
 DB 106 GAGGAAAT 114

RESULT 38

OY 09VZV2 PRELIMINARY; PRT; 696 AA.
 AC 09VZV2;
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE CG1869 protein.
 GN CG1869.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.

RC STRAIN=BERKELEY;
 RL MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galie R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotlier P.,
 RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,
 RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RT "The genome sequence of Drosophila melanogaster."
 RL Science 287:2185-2195(2000).
 DR EMBL: AE003477; AAF47714.1; -
 DR FLYbase: FBgn0035398; CG1869.
 DR InterPro: IPR001579; Chitinase_18/2.
 DR InterPro: IPR001223; Glyco_hydro_18.
 DR Pfam: PF00704; Glyco_hydro_18; 2.
 DR ProDom: PD000471; Glyco_hydro_18; 2.
 DR PROSITE: PS01095; CHITINASE_18; UNKNOWN_1.
 SQ SEQUENCE 696 AA; 77366 MW; 6877585837A19CF6 CRC64;

DR EMBL: AE002353; AAF39663.1; -.
 DR TIGR: TC0867; -.
 DR PRINTS: PRO1608; BACINVASINC.
 KW Hypothetical protein: Complete proteome.
 SQ SEQUENCE 491 AA; 50762 MW; 440F298930D93FB2 CRC64;

Query Match 1.8%; Score 9; DB 16; Length 491;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 447 FWOQASKIA 455
 DB 447 FWOQASKIA 455

RESULT 32

Q96LC4 PRELIMINARY; PRT; 502 AA.
 AC Q96LC4; 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Polyamine oxidase isoform-2.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wang Y., Devereux W., Stewart T.M., Casero R.A. Jr.;
 RT "Human polyamine oxidase, isoform-2."
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY033890; AAK5764.1; -.
 DR InterPro: IPR002937; Amino_oxidase.
 DR InterPro: IPR000205; NAD_binding.
 DR Pfam: PF01593; Amino_oxidase; 1.
 DR PRINTS: PRO0757; AMINEOXDASEF.
 SQ SEQUENCE 502 AA; 56156 MW; A6CA22481CFEED4 CRC64;

Query Match 1.8%; Score 9; DB 4; Length 502;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 GAGLAGLAA 288
 DB 31 GAGLAGLAA 39

RESULT 33

Q9BW38 PRELIMINARY; PRT; 502 AA.
 AC Q9BW38; 01-JUN-2001 (TREMBLrel. 17, Created)
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
 DE Similar to hypothetical protein.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Straussberg R.;
 RT Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: BC000669; AAH00669.1; -.
 DR InterPro: IPR000759; Adrxn_reductase.
 DR InterPro: IPR001613; Aminoacid fl.
 DR InterPro: IPR002937; Amino_oxidase.
 DR InterPro: IPR000205; NAD_binding.
 DR Pfam: PF01593; Amino_oxidase; 1.
 DR PRINTS: PRO0419; ADXRDYASE.
 DR PRINTS: PRO0757; AMINEOXDASEF.
 KW Hypothetical protein.

SQ SEQUENCE 502 AA; 56091 MW; 45921536D1B92403 CRC64;

Query Match 1.8%; Score 9; DB 4; Length 502;
 Best Local Similarity 100.0%; Pred. No. 13;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 GAGLAGLAA 288
 DB 31 GAGLAGLAA 39

RESULT 34

Q9NMW0 PRELIMINARY; PRT; 555 AA.
 AC Q9NMW0; 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
 DE CDNA FLJ20746 fis, clone HEP06040.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,
 RA Nakajima Y., Mizuno T., Morinaga M., Ota T., Suzuki Y., Obayashi M.,
 RA Nishi T., Shibahara T., Tanaka T., Nakamura Y., Isgal T., Sugano S.;
 RT "NEBO human cDNA sequencing project."
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AK000753; BAA91360.1; -.
 DR InterPro: IPR000759; Adrxn_reductase.
 DR InterPro: IPR002937; Amino_oxidase.
 DR InterPro: IPR000205; NAD_binding.
 DR Pfam: PF01593; Amino_oxidase; 1.
 DR PRINTS: PRO0419; ADXRDYASE.
 SQ SEQUENCE 555 AA; 61819 MW; BDBEA65CEB9FA5BF CRC64;

Query Match 1.8%; Score 9; DB 4; Length 555;
 Best Local Similarity 100.0%; Pred. No. 14;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 GAGLAGLAA 288
 DB 31 GAGLAGLAA 39

RESULT 35

Q96GT3 PRELIMINARY; PRT; 555 AA.
 AC Q96GT3; 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE Polyamine oxidase isoform-1.
 OS Homo sapiens (human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21347206; PubMed=11454677;
 RA Wang Y., Devereux W., Woster P.M., Stewart T.M., Hacker A.,
 RA Casero R.A. Jr.;
 RT "Cloning and characterization of a human polyamine oxidase that is
 inducible by polyamine analogue exposure."
 RT Cancer Res. 61:5370-5373(2001).
 RL EMBL: AY033889; AAK5763.1; -.
 DR InterPro: IPR002937; Amino_oxidase.
 DR InterPro: IPR000205; NAD_binding.
 DR Pfam: PF01593; Amino_oxidase; 1.
 SQ SEQUENCE 555 AA; 61872 MW; F0087D3629013C51 CRC64;

RESULT 28
 O9SF45 PRELIMINARY: PRT: 477 AA.
 ID 09SF45
 AC 09SF45
 DT 01-MAY-2000 (TREMBLrel. 13, Created)
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Putative oxidoreductase.
 GN F11F8.15 OR A3609580.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
 OC eustosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=CV. COLUMBIA;
 RA Lin X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,
 RA Maiti R., Romling C.M., Koo H., Fujii C.Y., Uterback T.R.,
 RA Barstead M.E., Bowman C.L., White O., Niernan W.C., Fraser C.M.;
 RT "Arabidopsis thaliana chromosome 3 BAC F11F8 genomic sequence.";
 RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
 RM [2]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Banno F., Chang E., Dale J.M.; Goldsmith A.D.,
 RA Lee J.M., Onodera C.S., Quach H.L., Tang C., Toriumi M., Wu H.C.,
 RA Yamamura Y., Yu G., Yu S., Bowser L., Carninci P., Chen H., Cheuk R.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Kim C., Kosena E., Lam B., Lin J., Meyers M.C., Miranda M.,
 RA Narasaka M., Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M.,
 RA Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RT "Full length cDNA of gene At3g09580 (GI:15232718)";
 RL Submitted (NOV-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AC016661; AAF23289.1; -;
 DR EMBL: AY064000; AAL36356.1; -;
 DR InterPro: IPR001613; Aminoacid.fl.
 DR InterPro: IPR002937; Amino.oxidase.
 DR Pfam: PF01593; Amino.oxidase; 1.
 DR PRINTS: PR00757; AMINOXDASEF.
 DR SEQUENCE 477 AA: 52258 MW: C332782957852BE9 CRC64;
 SQ
 Query Match 1.8%; Score 9; DB 10; Length 477;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 280 GAGLAGLAA 288
 Db 53 GAGLAGLAA 61

DR EMBL: AE001328; AAC68180.1; -;
 KM Hypothetical protein; Complete proteome.
 SQ SEQUENCE 487 AA: 50217 MW: DAD6038ECIDA03D CRC64;
 Query Match 1.8%; Score 9; DB 16; Length 487;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 447 FWOQASKIA 455
 Db 443 FWOQASKIA 451
 RESULT 30
 O9STX7 PRELIMINARY: PRT: 491 AA.
 ID 09STX7
 AC 09STX7
 DT 01-DEC-2001 (TREMBLrel. 19, Created)
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)
 DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
 DE GH25390P.
 GN ORK1 OR CG1615.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=BERKELEY;
 RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,
 RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,
 RA Gonzalez M., Guarin H., Li P., Liao G., Miranda A., Mungall C.J.,
 RA Nunoo J., Pacleb J., Paragas V., Park S., Phouanavong S., Wan K.,
 RA Yu C., Lewis S.E., Rubin G.M., Ceiniker S.;
 RT Submitted (OCT-2001) to the EMBL/Genbank/DBJ databases.
 DR EMBL: AY058455; AAL13684.1; -;
 DR Flybase: FBgn0017561; ORK1.
 DR SEQUENCE 491 AA: 50962 MW: 282FBE74A3E90506 CRC64;
 SQ
 Query Match 1.8%; Score 9; DB 5; Length 491;
 Best Local Similarity 100.0%; Pred. No. 12;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 294 AAAAGGAAG 302
 Db 254 AAAAGGAAG 262
 RESULT 31
 O9PJG2 PRELIMINARY: PRT: 491 AA.
 ID 09PJG2
 AC 09PJG2
 DT 01-OCT-2000 (TREMBLrel. 15, Created)
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
 DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
 DE Hypothetical protein TC0867.
 GN TC0867.
 OS Chlamydia muridarum.
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
 OX NCBI_TaxID=83560;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=MOPN / NIGG;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Uterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwin M., Nelson W., Deboy R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia pneumoniae AR39".
 RT Nucleic Acids Res. 28:1397-1406(2000).
 RL

```
DR InterPro: IPR000205; NAD_binding.
DR Pfam: PF01593; Amino_oxidase; 1.
DR PRINTS: PRO0419; ADXRDPTASE.
DR PROSITE: PS00225; CRYSTALLIN_BETAGAMMA; UNKNOWN_1.
DR NON_TER 285 285
SQ SEQUENCE 285 AA; 31164 MW; 3E7622A42FCCE35A CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 7.9;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 GAGLAGLAA 288
Db 7 GAGLAGLAA 15

RESULT 25
Q9XM22 PRELIMINARY; PRT; 307 AA.
AC Q9XM22;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE ADP/ATP translocase.
GN AAC.
OS Ascaris suum (Pig roundworm) (Ascaris lumbricoides).
OG Mitochondrion.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Ascaridida; Ascaridoidea;
OC Ascarididae; Ascaris.
OX NCBI_TaxID=6253;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=MUSCLE;
RA Harmyeh S.H., Komuniecki R.;
RT "Identification of a cDNA coding for ADP/ATP translocase in the
RT parasitic nematode Ascaris suum.";
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF130365; AAB30505.1;
DR InterPro: IPR001993; Mitoch_carrier.
DR InterPro: IPR002067; Mit_carrier.
DR Pfam: PF00153; mito_carr; 3.
DR PRINTS: PR00926; MITOCARRIER.
DR PROSITE: PS00215; MITOCH_CARRIER; UNKNOWN_2.
KW Mitochondrion.
SQ SEQUENCE 307 AA; 33575 MW; BD575E3AFAF21F3F5-CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 AAGGAAGAA 304
Db 132 AAGGAAGAA 140

RESULT 26
Q9NPY2 PRELIMINARY; PRT; 308 AA.
AC Q9NPY2;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE DJ779E11.1.1 (novel flavin containing amine oxidase (isoform 1)).
GN DJ779E11.1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Babbage A.;
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL121675; CAC00598.1; -.
```

```
DR InterPro: IPR000759; Adnrx_reductase.
DR InterPro: IPR002937; Amino_oxidase.
DR InterPro: IPR000960; Flav_cont_moxgn.
DR InterPro: IPR000205; NAD_binding.
DR Pfam: PF01593; Amino_oxidase; 1.
DR PRINTS: PRO0419; ADXRDPTASE.
DR PRINTS: PR00370; FMOXYGENASE.
SQ SEQUENCE 308 AA; 33734 MW; FCD4BF469B23D84F CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 8.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 GAGLAGLAA 288
Db 31 GAGLAGLAA 39

RESULT 27
Q8UAI1 PRELIMINARY; PRT; 474 AA.
AC Q8UAI1;
DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBlrel. 21, Last annotation update)
DE Secretion protein, HlyD family.
DE PRSE OR ATG3382 OR AGR L.2881.
GN Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OS Agrobacterium tumefaciens (strain C58 / ATCC 33970).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=176299;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608550; PubMed=11743193;
RA Wood D.W., Setubal J.C., Kaul R., Moks D.E., Kitajima J.P.,
RA Okura V.K., Zhou Y., Chen L., Wood G.E., Almeida N.F., Jr., Woo L.,
RA Chen Y., Paulsen I.T., Eisen J.A., Karp P.D., Boye D. Sr.,
RA Chapman P., Clendenning J., Deatherage G., Gillet M., Grant C.,
RA Kityavin T., Levy R., Li M.-J., McClelland E., Palmeri A.,
RA Raymond C., Rouse G., Saenphimachak C., Wu Z., Romero P., Gordon D.,
RA Zhang S., Yoo H., Tao Y., Biddle P., Jung M., Krespan W., Perry M.,
RA Gordon-Kamm B., Liao L., Kim S., Hendrick C., Zhao Z.-Y., Dolan M.,
RA Chumley F., Tingey S.V., Tomb J.-F., Gordon M.P., Olson M.V.,
RA Nester E.W.;
RT "The genome of the natural genetic engineer Agrobacterium tumefaciens
RT C58.";
RL Science 294:2317-2323(2001).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21608551; PubMed=11743194;
RA Goodner B., Hinkle G., Gallung S., Miller N., Blanchard M., Mullin L.,
RA Ourullo B., Goldman B.S., Cao Y., Askenazi M., Halling C., Mullin L.,
RA Houmiel K., Gordon J., Vaudin M., Iartchouk O., Epp A., Liu F.,
RA Wollam C., Allinger M., Doughty D., Scott C., Lappas C., Markez B.,
RA Flanagan C., Crowell C., Gursen J., Lomo C., Sear C., Strub G.,
RA Cielo C., Slater S.;
RT "Genome sequence of the plant pathogen and biotechnology agent
RT Agrobacterium tumefaciens C58.";
RL Science 294:2323-2328(2001).
DR EMBL: AE009268; AAL44195.1; ALT_INIT.
DR EMBL: AE008344; AAK90008.1; -.
KW Complete proteome.
SQ SEQUENCE 474 AA; 52661 MW; 8ED17141B786EAC7 CRC64;

Query Match
Best Local Similarity 100.0%; Pred. No. 12;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 283 IAGLAGAA 291
Db 56 IAGLAGAA 64
```



```

RA Kormanec J., Homeroova D., Sevcikova B., Novakova R., Blaskovic D.;
RT "Large subunit of glutamate synthase of Streptomyces coelicolor
RL A3(2)";
RL Submitted (JUN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF162938; AAD49328.1; -.
DR InterPro; IPR002839; DUF125.
DR Pfam; PF01988; DUF125; 1.
FT NON-TER
SQ SEQUENCE 239 AA; 24777 MW; 5559563D7E5E962C CRC64;

Query Match
Best Local Similarity 1.8%; Score 9; DB 2; Length 239;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 282 GLAGLAAGA 290
Db 59 GLAGLAAGA 67

RESULT 22
O9S2Y8 PRELIMINARY; PRT; 243 AA.
AC O9S2Y8;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative membrane protein.
GN SC02027 OR SC3A3.05c.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RT SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RT SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA James K.D., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (AUG-1999) to the EMBL/GenBank/DBJ databases.
RN [3]
RT SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,
RA Kinashi H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL Mol. Microbiol. 21:77-96(1996).
RN [4]
RT SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Latke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)";
RL Nature 417:141-147(2002).
DR EMBL; AL109849; CAB52862.1; -.
DR InterPro; IPR002839; DUF125.
DR Pfam; PF01988; DUF125; 1.
SQ SEQUENCE 243 AA; 25205 MW; B3AF946838905DE1 CRC64;

Query Match
Best Local Similarity 1.8%; Score 9; DB 16; Length 243;
Matches 9; Conservative 100.0%; Pred. No. 6.9;

```

```

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 282 GLAGLAAGA 290
Db 63 GLAGLAAGA 71

RESULT 23
O9K3G4 PRELIMINARY; PRT; 247 AA.
AC O9K3G4;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Hypothetical protein SC01277.
GN SC01277 OR 2SCG18.24.
OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RT SEQUENCE FROM N.A.
RC STRAIN=A3(2) / M145;
RA Bentley S.D., Chater K.F., Cerdeno-Tarraga A.-M., Challis G.L.,
RA Thomson N.R., James K.D., Harris D.E., Quail M.A., Kieser H.,
RA Harper D., Bateman A., Brown S., Chandra G., Chen C.W., Collins M.,
RA Cronin A., Fraser A., Goble A., Hidalgo J., Hornsby T., Howarth S.,
RA Huang C.-H., Kieser T., Latke L., Murphy L., Oliver K., O'Neill S.,
RA Rabinowitsch E., Rajandream M.A., Rutherford K., Rutter S.,
RA Seeger K., Saunders D., Sharp S., Squares R., Squares S., Taylor K.,
RA Warren T., Wietzorrek A., Woodward J., Barrell B.G., Parkhill J.,
RA Hopwood D.A.;
RT "Complete genome sequence of the model actinomycete Streptomyces
RT coelicolor A3(2)";
RL Nature 417:141-147(2002).
DR EMBL; AL390188; CAB99157.1; -.
KW Hypothetical protein.
SQ SEQUENCE 247 AA; 22950 MW; EE32ABD76146881B CRC64;

Query Match
Best Local Similarity 1.8%; Score 9; DB 16; Length 247;
Matches 9; Conservative 100.0%; Pred. No. 7;

Oy 283 LAGLAAGA 291
Db 96 LAGLAAGA 104

RESULT 24
O9RCT2 PRELIMINARY; PRT; 285 AA.
AC O9RCT2;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Zeta carotene desaturase (Fragment).
GN CRTQ.
OS Synecococcus leopoliensis.
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.
OX NCBI_TaxID=32047;
RN [1]
RT SEQUENCE FROM N.A.
RC STRAIN=SAUG 1402-1;
RA MEDLINE=20465103; PubMed=11007968;
RA Miller B., Heuser T., Zimmer W.;
RT "Functional involvement of a deoxy-D-xylose 5-phosphate
RT reductoisomerase gene harboring locus of Synecococcus leopoliensis in
RT isoprenoid biosynthesis.";
RL FEBS Lett. 481:221-226(2000).
DR EMBL; AJ250721; CAB65434.1; -.
DR InterPro; IPR000759; Adrxn_reductase.
DR InterPro; IPR001064; Crystallin.

```

```

DF 01-NOV-1996 (TREMBLrel. 01, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE Peripheral benzodiazepine receptor.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=PLACENTA;
RX MEDLINE=94140364; PubMed=8307574;
RA Lin D., Chang Y.J., Straus J.F., Miller W.L.;
RT "The human peripheral benzodiazepine receptor gene: cloning and
RT characterization of alternative splicing in normal tissues and in a
RT patient with congenital lipid adrenal hyperplasia.";
RL Genomics 18:643-650(1993).
DR EMBL; L21954; AAA18228.1; -.
DR EMBL; L21952; AAA18228.1; JOINED.
DR EMBL; L21953; AAA18228.1; JOINED.
DR InterPro: IPR004307; Tspo_MBR.
DR Pfam: PF03073; Tspo_MBR; 1.
KW Receptor.
SQ SEQUENCE 169 AA; 18483 MW; C353A23AFADF327E CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 4; Length 169;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 GAAATTTVA 310
DB 117 GAAATTTVA 125
|||||
RESULT 18
ID Q96TF6 PRELIMINARY; PRT; 169 AA.
AC Q96TF6;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, last annotation update)
DE Df326114.1 (benzodiazepine receptor (peripheral) (PBR, PKBS,
DE mitochondrial benzodiazepine, MBR) (isoform 1)).
GN BZRP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Badbage A.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; Z82214; CAB55884.1; -.
DR InterPro: IPR004307; Tspo_MBR.
DR Pfam: PF03073; Tspo_MBR; 1.
SQ SEQUENCE 169 AA; 18828 MW; 6DD741BF99AB89D9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 4; Length 169;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 302 GAAATTTVA 310
DB 117 GAAATTTVA 125
|||||
RESULT 19
ID Q9URS5 PRELIMINARY; PRT; 178 AA.
AC Q9URS5;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE CRN1 homologue (Fragment).

```

```

GN CRN1.
OS Kluyveromyces lactis (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Kluyveromyces.
OX NCBI_TaxID=28985;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2359/152; PubMed=10669871;
RX MEDLINE=20136097;
RA Bao W.G., Huo K.K., Li Y.Y., Fukuhara H.;
RT "Protein disulfide isomerase genes of Kluyveromyces lactis.";
RL Yeast 16:329-341(2000).
DR EMBL; AJ243960; CAB51777.1; -.
FT NON_TER
SQ SEQUENCE 178 AA; 18698 MW; C3D8E079CE461A CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 3; Length 178;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 298 GGAGAGAAA 306
DB 106 GGAGAGAAA 114
|||||
RESULT 20
ID Q96LC3 PRELIMINARY; PRT; 190 AA.
AC Q96LC3;
DT 01-DEC-2001 (TREMBLrel. 19, Created)
DT 01-DEC-2001 (TREMBLrel. 19, last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, last annotation update)
DE Polyamine oxidase isoform-3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang Y., Deyverux W., Stewart T.M., Hacker A., Casero R.A. Jr.;
RT "Human polyamine oxidase, isoform-3.";
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY033891; AAK55765.1; -.
DR InterPro: IPR002937; Amino_oxidase.
DR InterPro: IPR002005; NAD_binding.
DR Pfam: PF01593; Amino_oxidase; 1.
SQ SEQUENCE 190 AA; 20616 MW; DE0614286373A56A CRC64;

Query Match
Best Local Similarity 100.0%; Score 9; DB 4; Length 190;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 280 GAGLAGLAA 288
DB 31 GAGLAGLAA 39
|||||
RESULT 21
ID Q9S440 PRELIMINARY; PRT; 239 AA.
AC Q9S440;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)
DE H3U (Fragment).
GN H3U.
OS Streptomyces coelicolor A3(2).
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=100226;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=M145;

```

```

ID 094F07 PRELIMINARY; PRT; 147 AA.
AC 094F07;
AC 094F02:
AC 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Glycine-rich protein GRP20.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=21324873; PubMed=11431566;
RX Mayfield J.A., Fiebig A., Johnstone S.E., Preuss D.;
RT "Gene families from the Arabidopsis thaliana pollen coat proteome.";
RL Science 292:2482-2485(2001).
DR EMBL; AF362475; AAK83832.1; -
DR InterPro: IPR000136; Oleosin.
DR Pfam: PF01277; Oleosin; 1.
SQ SEQUENCE 147 AA; 14878 MW; C3FEE84BC4F5C0F0 CRC64;

Query Match
Best Local Similarity 1.8%; Score 9; DB 10; Length 147;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGAA 301
DB 119 GAAAGGAA 127

RESULT 14
094F02 PRELIMINARY; PRT; 149 AA.
AC 094F02;
AC 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Glycine-rich protein GRP20.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=21324873; PubMed=11431566;
RX Mayfield J.A., Fiebig A., Johnstone S.E., Preuss D.;
RT "Gene families from the Arabidopsis thaliana pollen coat proteome.";
RL Science 292:2482-2485(2001).
DR EMBL; AF362476; AAK83838.1; -
DR InterPro: IPR000136; Oleosin.
DR Pfam: PF01277; Oleosin; 1.
SQ SEQUENCE 149 AA; 15091 MW; FF5B139BAA2D84E8 CRC64;

Query Match
Best Local Similarity 1.8%; Score 9; DB 10; Length 149;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGAA 301
DB 119 GAAAGGAA 127

RESULT 15
094F07 PRELIMINARY; PRT; 153 AA.
AC 094F07;
AC 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE Oleosin-like protein (glycine-rich protein GRP20).
GN T211_270.

```

```

OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RP Beyer M., Murphy G., Ridley P., Hudson S., Bancroft I., Mewes H.W.,
RA Rudd S., Lemcke K., Mayer K.F.X.;
RA Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
RL [2]
RP SEQUENCE FROM N.A.
RP EU Arabidopsis sequencing project;
RA Submitted (Apr-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RP MEDLINE=21324873; PubMed=11431566;
RX Mayfield J.A., Fiebig A., Johnstone S.E., Preuss D.;
RT "Gene families from the Arabidopsis thaliana pollen coat proteome.";
RL Science 292:2482-2485(2001).
DR EMBL; AL163912; CAB87945.1; -
DR EMBL; AF362477; AAK83839.1; -
DR InterPro: IPR000136; Oleosin.
DR Pfam: PF01277; Oleosin; 1.
SQ SEQUENCE 153 AA; 15458 MW; 8D04492C2F5EB39B CRC64;

Query Match
Best Local Similarity 1.8%; Score 9; DB 10; Length 153;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGAA 301
DB 119 GAAAGGAA 127

RESULT 16
094F02 PRELIMINARY; PRT; 153 AA.
AC 094F02;
AC 01-DEC-2001 (TREMBlrel. 19, Created)
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)
DE Glycine-rich protein GRP20.
OS Arabidopsis thaliana (Mouse-ear cross).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RP MEDLINE=21324873; PubMed=11431566;
RX Mayfield J.A., Fiebig A., Johnstone S.E., Preuss D.;
RT "Gene families from the Arabidopsis thaliana pollen coat proteome.";
RL Science 292:2482-2485(2001).
DR EMBL; AF362474; AAK83826.1; -
DR InterPro: IPR000136; Oleosin.
DR Pfam: PF01277; Oleosin; 1.
SQ SEQUENCE 153 AA; 15486 MW; 8D04492C3DD5B39B CRC64;

Query Match
Best Local Similarity 1.8%; Score 9; DB 10; Length 153;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGAA 301
DB 119 GAAAGGAA 127

RESULT 17
013850 PRELIMINARY; PRT; 169 AA.
ID 013850;
AC 013850;
AC 01-NOV-1996 (TREMBlrel. 01, Created)

```

Query Match 1.8%; Score 9; DB 3; Length 106;
 Best Local Similarity 100.0%; Pred. No. 3.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAAAGGAG 302
 |||||||||
 DB 67 AAAAGGAG 75

RESULT 9

O8TFM9 PRELIMINARY: PRT: 109 AA.
 AC O8TFM9:
 DT 01-JUN-2002 (TREMblrel. 21, Created)
 DT 01-JUN-2002 (TREMblrel. 21, last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
 DE 60S acidic ribosomal protein P2.
 OS Fusarium culmorum.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Hypocreales; mitosporic Hypocreales; Fusarium.
 OX NCBI_TaxID=5516;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Hoff M., Krall M., Vieths S., Haustein D.;
 RT "IGE-binding acidic ribosomal protein P2 is a potential allergen in
 RT Fusarium culmorum."
 RL Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AY077706; AAL79930.1;
 KM Ribosomal protein.
 SQ SEQUENCE 109 AA; 11036 MW; D20E40DB073BDF77 CRC64;

Query Match 1.8%; Score 9; DB 3; Length 109;
 Best Local Similarity 100.0%; Pred. No. 3.5;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGGAA 301
 |||||||||
 DB 76 GAAAGGAA 84

RESULT 10

O9C3Z5 PRELIMINARY: PRT: 111 AA.
 AC O9C3Z5:
 DT 01-JUN-2001 (TREMblrel. 17, Created)
 DT 01-JUN-2001 (TREMblrel. 17, last sequence update)
 DT 01-OCT-2001 (TREMblrel. 18, last annotation update)
 DE Ribosomal protein P2.
 OS Podospora anserina.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariales; Lasiosphaeriaceae; Podospora.
 OX NCBI_TaxID=5145;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Vierny C., Folichon M., Silar P.;
 RT "Ribosomal proteins of Podospora anserina."
 RL Submitted (DEC-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF331715; AKK11263.1;
 DR InterPro: IPR001813; 60S-ribosomal.
 DR InterPro: IPR001859; Ribosomal.P2.
 DR Pfam: PF00428; 60S-ribosomal.1.
 DR PRINTS: PR00456; RIBOSOMALP2.
 SQ SEQUENCE 111 AA; 10975 MW; 94D7379F9E73C414 CRC64;

Query Match 1.8%; Score 9; DB 3; Length 111;
 Best Local Similarity 100.0%; Pred. No. 3.6;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGGAA 301
 |||||||||
 DB 72 GAAAGGAA 80

RESULT 11

ID O8V7C9 PRELIMINARY: PRT: 128 AA.
 AC O8V7C9:
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, last sequence update)
 DT 01-JUN-2002 (TREMblrel. 21, last annotation update)
 DE ORF2.
 OS TT virus.
 OC Viruses; ssDNA viruses; unclassified ssDNA viruses.
 OX NCBI_TaxID=68887;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA MEDLINE=21844401; PubMed=11855633;
 RA Peng Y.H., Nishizawa T., Takahashi M., Ishikawa T., Yoshikawa A.,
 RA Okamoto H.;
 RT "Analysis of the complete genomes of thirteen TT virus variants
 RT classifiable into the fourth and fifth genetic groups, isolated from
 RT viremic infants."
 RL Arch. Virol. 147:21-41(2002).
 DR EMBL: AB064624; BAB79391.1;
 DR InterPro: IPR004118; TT-ORF2.
 DR Pfam: PF02957; TT-ORF2.1.
 SQ SEQUENCE 128 AA; 13219 MW; 114D2C5DF694C37 CRC64;

Query Match 1.8%; Score 9; DB 12; Length 128;
 Best Local Similarity 100.0%; Pred. No. 4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 298 GGAGAGAAA 306
 |||||||||
 DB 97 GGAGAGAAA 105

RESULT 12

ID O76068 PRELIMINARY: PRT: 143 AA.
 AC O76068:
 DT 01-NOV-1998 (TREMblrel. 08, Created)
 DT 01-NOV-1998 (TREMblrel. 08, last sequence update)
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)
 DE Peripheral-type benzodiazepine receptor (Fragment).
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=BREAST;
 RA Hardwick M.J., Papadopoulos V.;
 RT "Peripheral-type benzodiazepine Receptor (PBR) in Human Breast
 RT Cancer."
 RT Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
 DR EMBL: AF075590; AAC31173.1;
 DR EMBL: AF075589; AAC31172.1;
 DR InterPro: IPR004307; Tspo_MBR.
 DR Pfam: PF03073; Tspo_MBR.1.
 FT NON_TER
 SQ SEQUENCE 143 AA; 16107 MW; DEAF74A8415C65A9 CRC64;

Query Match 1.8%; Score 9; DB 4; Length 143;
 Best Local Similarity 100.0%; Pred. No. 4.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 302 GAAATTTVA 310
 |||||||||
 DB 91 GAAATTTVA 99

RESULT 13
 O94F07

RA Moss B.;
RL Submitted (JUN-1996) to the EMBL/Genbank/DBJ databases.
DR EMBL: U60315; AAC55240.1; -;
SQ SEQUENCE 128 AA; 12728 MW; 22334EDB0B724607 CRC64;

Query Match 2.0%; Score 10; DB 12; Length 128;
Best Local Similarity 100.0%; Pred. No. 0.46;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 287 AAGAAGVAAA 296
DB 83 AAGAAGVAAA 92

RESULT 5

ID 036896 PRELIMINARY; PRT; 137 AA.
AC 036896;
DT 01-JAN-1998 (TREMBLrel. 05, Created)
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE GAG polyprotein (Fragment).
GN GAG.
OS Human immunodeficiency virus type 1.
OC Viruses; Retroid viruses; Retroviridae; Lentivirus.
OX NCBI_TaxID=11676;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97445059; PubMed=9300048;
RA Leigh Brown A.J., Lobidel D., Wade C.M., Rebus S., Phillips N.,
RA Brettle R.P., France A.J., Leen C.S., McMenamin J., McMillan A.,
RA Maw R.D., Mulcahy F., Robertson J.R., Sankar K.N., Scott G., Wyld R.,
RA Peutherer J.F.;
RT "The molecular epidemiology of human immunodeficiency virus type 1 in
RT six cities in Britain and Ireland.";
RL Virology 235:166-177(1997).
DR EMBL: AF014293; AAC58374.1; -;
DR InterPro: IPR000071; Retrovir_p17.
DR Pfam: PF00540; Gag_p17; 1.
KW AIDS; Core protein; Polyprotein.
FT NON_TER 1
FT NON_TER 137
SQ SEQUENCE 137 AA; 14958 MW; 409CDC38A4F61522 CRC64;

Query Match 2.0%; Score 10; DB 15; Length 137;
Best Local Similarity 100.0%; Pred. No. 0.49;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 287 AAGAAGVAAA 296
DB 96 AAGAAGVAAA 105

RESULT 6

ID 09FU65 PRELIMINARY; PRT; 487 AA.
AC 09FU65;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE Putative non-phototropic hypocotyl 3.
GN P0013F10.9.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Euphorbiaceae; Oryzaeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0013F10.";

RL Submitted (JUN-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AP002523; BAB17063.1; -;
DR InterPro: IPR004249; NPH3.
DR Pfam: PF03000; NPH3; 1.
SQ SEQUENCE 487 AA; 50508 MW; 84B57B75E64AFC11 CRC64;

Query Match 2.0%; Score 10; DB 10; Length 487;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 298 GGAAGAAAT 307
DB 468 GGAAGAAAT 477

RESULT 7

ID 08S076 PRELIMINARY; PRT; 103 AA.
AC 08S076;
DT 01-JUN-2002 (TREMBLrel. 21, Created)
DT 01-JUN-2002 (TREMBLrel. 21, Last sequence update)
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)
DE P0470A12.43 protein.
GN P0470A12.43.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Euphorbiaceae; Oryzaeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. NIPPONBARE;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
RT clone:P0470A12.";
RL Submitted (MAR-2001) to the EMBL/Genbank/DBJ databases.
DR EMBL: AP003436; BAB90315.1; -;
SQ SEQUENCE 103 AA; 10426 MW; 4A1EA91B1477B35B CRC64;

Query Match 1.8%; Score 9; DB 10; Length 103;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 294 AAAGGAAG 302
DB 17 AAAGGAAG 25

RESULT 8

ID 09HF07 PRELIMINARY; PRT; 106 AA.
AC 09HF07;
DT 01-MAR-2001 (TREMBLrel. 16, Created)
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE 60S acidic ribosomal protein type P1-A.
GN P1A.
OS Candida albicans (Yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.
OX NCBI_TaxID=5476;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CCM 8215;
RA Abramczyk D., Tchorzewski M., Grankowski N.;
RT "Cloning, expression and purification of the acidic ribosomal protein
RT from Candida albicans.";
RL Submitted (NOV-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL: AF317659; AAG33240.1; -;
DR InterPro: IPR001813; 60S_ribosomal.
DR Pfam: PF00428; 60S_ribosomal; 1.
KW Ribosomal protein.
SQ SEQUENCE 106 AA; 10996 MW; BB03BB1B0CE9547E CRC64;

Query Match	98.2%;	Score 487;	DB 16;	Length 488;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 487;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps
QY 4	MSISSSSGPNOKNINISOVLTPTPGVPOODL	LSGNETTKOIQOOTROGKNTMESD	ATAG 63	
Db 1	MSISSSSGPNOKNINISOVLTPTPGVPOODL	LSGNETTKOIQOOTROGKNTMESD	ATAG 60	
QY 64	ASGDKDTSSTTKETAPQGVAAAGKSSS	QKAGADTVGSAAATSNFTTKIAMQTSI	123	
Db 61	ASGDKDTSSTTKETAPQGVAAAGKSSS	QKAGADTVGSAAATSNFTTKIAMQTSI	120	
QY 124	EEAKSMESTLESISQISAQOMEVEAVVVAAL	SGKSSSAAKLETPPELPKPGVTPRSEVI	183	
Db 121	EEAKSMESTLESISQISAQOMEVEAVVVAAL	SGKSSSAAKLETPPELPKPGVTPRSEVI	180	
QY 184	EIGLALAKAIOTTGAEATKSAISNAST	QAOADPTNKLGEKQAIKIDKRETEYQEMKAE	243	
Db 181	EIGLALAKAIOTTGAEATKSAISNAST	QAOADPTNKLGEKQAIKIDKRETEYQEMKAE	240	
QY 244	QSKSLEGTMDTVMTVMIAVSAITTVISIAAI	FTGCGAGLAGIAGAAGAAAGAGAGA 303		
Db 241	QSKSLEGTMDTVMTVMIAVSAITTVISIAAI	FTGCGAGLAGIAGAAGAAAGAGAGA 300		
QY 304	AAATTVAOTITVOAVVQAVKQAVTTANRQAI	TTAAIKRAVSGIKATPKTLVYIAKAISK 363		
Db 301	AAATTVAOTITVOAVVQAVKQAVTTANRQAI	TTAAIKRAVSGIKATPKTLVYIAKAISK 360		
QY 364	GISKVFAKGTQMIKAFNPKLSKVISISLTSK	WTVGVVVAAPALGKIMQOLSEMOCN 423		
Db 361	GISKVFAKGTQMIKAFNPKLSKVISISLTSK	WTVGVVVAAPALGKIMQOLSEMOCN 420		
QY 424	VAPQKEVKGKLOAAADMTSMFTQPMQQA	SKIASKQTESNEMQKAKLKLOALIKRAYAI 483		
Db 421	VAPQKEVKGKLOAAADMTSMFTQPMQQA	SKIASKQTESNEMQKAKLKLOALIKRAYAI 480		
QY 484	SGAIGA 490			
Db 481	SGAIGA 487			
RESULT 3				
ID 09W3L3	PRELIMINARY;	PRT: 1937 AA.		
AC 09W3L3:				
DT 01-MAY-2000 (TREMBLrel. 13, Created)				
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)				
DT 01-JUN-2002 (TREMBLrel. 21, Last annotation update)				
DE Female sterile (1) homoetic protein.				
GN FS(1)H OR CG2252.				
OS Drosophila melanogaster (Fruit fly).				
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;				
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
OC Ephydroidea; Drosophilidae; Drosophila.				
ON NCBI_TaxID=7227;				
RA 11				
RP SEQUENCE FROM N.A.				
RC STRAIN=BERKELEY;				
RX MEDLINE=20196006; PubMed=10731132;				
RA Adams M.D., Celihner S.E., Holt R.A., Evans C.A., Gocayne J.D.,				
RA Amaratilake P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,				
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,				
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,				
RA Brandon R.C., Rogers J.-H.C., Baxter E.G., Hejt G., Nelson C.R., Miklos G.L.G.,				
RA Wan K.H., Doyle C., Baxter E.G., Hejt G., Nelson C.R., Miklos G.L.G.,				
RA Abril J.F., Abayaratne A., An H.-D., Andrews-Plattko C., Baldwin D.,				
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,				
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,				
RA Botkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,				
RA Butris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,				
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,				
RA de Pablos B., Detcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,				

	Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.
	RA Durbin R.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
	RA Glødestad A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
	RA Hariris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck C.,
	RA Hostall D., Houston K.A., Howland T.J., Wei M.-H., Ideyram C.,
	RA Joslin M., Kalush F., Kapren G.H., Ke Z., Kennison J.A., Ketchum K.A.,
	RA Kissel B.E., Kodira C.D., Kraft C., Kravitz S., Kulip D., Lai X.,
	RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
	RA Liu X., Mattei B., McIntosh T.C., Meleod M.P., Moherston D.,
	RA Mekullov G., Mishina N.V., Mobarry C., Morris J., Mostreil A.,
	RA Mount S.M., Moy M., Murphy B., Murphy L., Munzy D.M., Nelson D.L.,
	RA Nelsson D.R., Nielson K.A., Nixon K., Nusseron D.R., Paeblo J.M.,
	RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
	RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
	RA Shee B.C., Sidén-Klamos I., Simpson M., Skupski M.P., Smith T.,
	RA Splier E., Spreading A.C., Turner R., Venster E., Wang A.H., Wang X.,
	RA Syrkstas R., Tecor C., Turner R., Weinstock G.M., Weissbach J.,
	RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Wu D., Yang S., Yao Q.A.,
	RA Williams S., Woodgett T., Worley K.C., Wu D., Zhang G., Zhao Q., Zheng L.,
	RA Ye J., Yen R.-F., Zavrel J.S., Zhan M., Zhang X., Zhu S., Zhu X., Smith H.O.,
	RA Zhong X.H., Zhong F.N., Rubin G.M., Venster J.C.;
RA	Gibbs R.A., Myers E.W., Rubin G.M., Venster J.C.);
RT	"The genome sequence of <i>Drosophila melanogaster</i> ".;
RL	Science 287:2185-2195(2000).
DR	HSSP; AE003442; AAFA6312.2; -.
DR	HMBP; O92831; IB91.
DR	Ilybase: FBgn0004656; fs(1)h.
DR	Interpro: IPR001487; Bromodomaiin.
DR	Interpro: IPR002395; Kinloggen.
DR	Interpro: IPR002173; PFkB.
DR	Pfam: PF00439; bromodomaiin_2.
DR	PRINTS: PRO0303; BROMODOMAIN.
DR	PRINTS: PR00354; KININOSEN.
DR	SMART; SM00297; BKOMO; 2.
DR	PROSITE; PS00633; BROMODOMAIIN.1; 2.
DR	PROSITE; PS0014; BROMODOMAIIN.2; 2.
DR	PROSITE; PS00583; PFKB_KINASES.1; UNKNOWN.1.
SQ	SEQUENCE 1937 AA: 195339 MW: IDB0AAVBB3J1f06B CRC64:

QY	294	AAAAGGAGAA	304
Db	343	AAAAGGAGAA	353

RESULT 4	Q98279	PRELIMINARY;	PRT;	128 AA.
ID	Q98279			
AC	Q98279;			
DC	01-FEB-1997 (TrEMBLrel. 02, Created)			
DT	01-FEB-1997 (TrEMBLrel. 02, Last sequence update)			
DT	01-NOV-1998 (TrEMBLrel. 08, Last annotation update)			
DE	MC112L.			
NC	MC112L.			
OS	Molluscum contagiosum virus subtype 1 (MCV1).			
OC	Viruses; dsDNA viruses, no RNA stage; Poxviridae; Chordopoxvirinae;			
OC	Molluscipoxvirus.			
OX	NCBI_TaxID=10280;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96325459; PubMed=8670425;			
RA	Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,			
RA	Moss B.;			
RT	"Genome sequence of a human tumorigenic poxvirus: Prediction of			
RT	specific host response-evasion genes.";			
RL	Science 273:813-816(1996).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RA	Senkevich T.G., Bugert J.J., Sisler J.R., Koonin E.V., Darai G.,			

966	1.4	547	16	08U913	08u913 agrobacteri
967	1.4	548	10	08M378	08m378 oryza sativ
968	7	549	12	0918h6	0918h6 baboon cyto
969	7	549	17	097UN3	097un3 sulfolobus
970	7	551	10	09EV45	09efv45 tagetes ere
971	7	551	16	092SM7	092sm7 rhizobium m
972	7	552	10	094KE8	094ke8 citrus para
973	7	552	16	098C08	098c08 rhizobium l
974	7	553	5	09wV10	09wv10 drosophila
975	7	553	10	09ZF26	09zf16 citrus sine
976	7	554	16	09RFS0	09rfs0 delinococcus
977	7	555	16	09L022	09l022 streptomyces
978	7	557	16	09K410	09k410 streptomyces
979	7	559	17	09HRT5	09hrt5 halobacteri
980	7	560	16	09HYA3	09hya3 pseudomonas
981	7	563	5	09wM12	09wm12 drosophila
982	7	564	16	09AAY7	09aay7 caulobacter
983	7	565	3	01J510	01j510 rhodotorula
984	7	566	1	P77943	P77943 sulfolobus
985	7	566	10	093ZB4	093zb4 arabidopsis
986	7	566	16	P71838	P71838 mycobacteri
987	7	566	17	09C4L9	09c4l9 sulfolobus
988	7	568	5	09V600	09v600 drosophila
989	7	570	10	065813	065813 haematococcc
990	7	570	16	P84544	P84544 bacillus su
991	7	571	5	095047	095047 drosophila
992	7	572	10	022056	022056 arabidopsis
993	7	573	10	023212	023212 arabidopsis
994	7	574	4	09H7K5	09h7k5 homo sapien
995	7	574	10	022455	022455 arabidopsis
996	7	574	12	065563	065563 bovine herpi
997	7	575	16	091642	091642 pseudomonas
998	7	578	5	09wK24	09wk24 drosophila
999	7	578	10	093V74	093v74 oryza sativ
1000	7	580	10	08vX08	08vx08 gentiana lu

ALIGNMENTS

RESULT 1	
Q9Z797	
ID Q9Z797	PRELIMINARY; PRT; 493 AA

DT 01-MAY-1999 (TREMBLrel. 10, Created)
DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)
DT 01-MAR-2002 (TREMBLrel. 20, Last annotation update)
DE C1578 hypothetical protein.
GN C1578 hypothetical protein.
OS Chlamydia pneumoniae (Chlamydia phila pneumoniae).
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia phila.
OX NCBI_TaxId=83558;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C1029;
RX MEDLINE=9206606; PubMed=10192388;
RA Kaiman S., Mitchell W., Marathe R., Lammel C., Fan J., Hyman R.W., Davis R., Grimwood J., Matthews R.S., Stephens R.S.,
RT Comparative genomes of Chlamydia pneumoniae and C. trachomatis." (1999)
RL Nat. Genet. 21:385-389(1999)
DR EMBL AF001662; AAD18947.1; -
DR InterPro IPR00508; Slippase
DR PROSITE PS005017 SPASE_1.1; UNKNOWN_1.
KW Complete proteome.
SQ SEQUENCE 493 AA; 50489 MW; 90F315B7A20AC2F0 CRC64;

Query Match	99.48;	Score 493;	DB 16;	Length 493;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 493;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

```

0y      4 MSISSSGPDNOKNIMSVLTSTPQVPPQDKLSGNETKQIQOTRQKNTMESDATTAG 63
        |||||||
        1 MSISSSGPDNOKNIMSVLTSTPQVPPQDKLSGNETKQIQOTRQKNTMESDATTAG 60
Db

```

QY	64	ASGDKDTSMTTKTEPTAAQOOGVAAGKESSESOKAGADTVSGAAATTAASNTATKIAMOTSI	123
Dp	61	ASGDKDTSMTTKTEPTAAQOOGVAAGKESSESOKAGADTVSGAAATTAASNTATKIAMOTSI	120
QY	124	EEAKSSEMTLESLOSLSAOMKEVEAVVVAALSGKSSGSAKLETPELPRGYTPREVI	183
Dp	121	EEAKSSEMTLESLOSLSAOMKEVEAVVVAALSGKSSGSAKLETPELPRGYTPREVI	180
QY	184	ETGLALAKAIOTLCEATKRSALSNASTAOAADOTNKGLEKOAIKIDKEEVEYOEKKAAC	243
Dp	181	ETGLALAKAIOTLCEATKRSALSNASTAOAADOTNKGLEKOAIKIDKEEVEYOEKKAAC	240
QY	244	OKSKDLEGTMDTVMTWMIANVASVATTVSIYAAIPTCGAGLAGLAGAAGVAAAAAGAGA	303
Dp	241	OKSKDLEGTMDTVMTWMIANVASVATTVSIYAAIPTCGAGLAGLAGAAGVAAAAAGAGA	300
QY	304	AAATTVTAQTITQAVVOAVKQAVITTAAROAITAAIKAAVKSIGIAFIKTLYKAIKAIK	363
Dp	301	AAATTVTAQTITQAVVOAVKQAVITTAAROAITAAIKAAVKSIGIAFIKTLYKAIKAIK	360
QY	364	GISVPEFKGTOMIKNPKKISKVSIISLTSMWTVGVVVVAAPLKGIMOMOLESEON	423
Dp	361	GISVPEFKGTOMIKNPKKISKVSIISLTSMWTVGVVVVAAPLKGIMOMOLESEON	420
QY	424	VAFQOKEVGKLOAAADIMISMTQFQWOQASKASKQTESNEMTQKARKLGAQILKAAAI	483
Dp	421	VAFQOKEVGKLOAAADIMISMTQFQWOQASKASKQTESNEMTQKARKLGAQILKAAAI	480
QY	484	SGAATAGAKHTNNF	496
Dp	481	SGAATAGAKHTNNF	493

RESULT 2

ID Q9JS66 PRELIMINARY; PRT; 488 AA.

DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last annotation update)
 DE C1578 hypothetical protein.
 GN CP00809 OR CP1062.
 OS Chlamydia pneumoniae (Chlamydia phila pneumoniae).
 OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia phila.
 OX NCBI_TaxID=83558;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=AR39;
 RX MEDLINE=20150255; PubMed=10684935;
 RA Read T.D., Brunham R.C., Shen C., Gill S.R., Heidelberg J.F.,
 RA White O., Hickey E.K., Peterson J., Uitterback T., Berry K., Bass S.,
 RA Linher K., Weidman J., Khouri H., Craven B., Bowman C., Dodson R.,
 RA Gwynn M., Neilson W., Debay R., Kolonay J., McClarty G., Salzberg S.L.,
 RA Eisen J., Fraser C.M.;
 RT "Genome sequences of Chlamydia trachomatis Mopn and Chlamydia
 RT pneumoniae AR39.";
 RL Nucleic Acids Res. 28:1397-1406(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=J138;
 RX MEDLINE=20330349; PubMed=10671362;
 RA Shiba T., Hirakawa H., Kimoto M., Tabuchi M., Kishi F., Ouchi K.,
 RA Shiba T., Ishii K., Hattori M., Kunara S., Nakazawa T.;
 RT "Comparison of whole genome sequences of Chlamydia pneumoniae J138
 RT from Japan and CWL029 from USA."
 RL Nucleic Acids Res. 28:2311-2314(2000).
 DR EMBL: AE002263; AAF38835.1; -;
 DR EMBL: AP002548; BAA99017.1; -;
 DR TIGR: CP1062; -;
 DR InterPro: IPR000508; SigPase.
 DR PROSITE: PS00501; SPASE_I.1; UNKNOWN_I.
 SO SEQUENCE 488 AA; 49819 MW; 06CF0B905398AAB CRC64;

820	7	1.4	444	5	Q24473	Q24473 drosophila	893	7	1.4	496	16	Q91692	Q91692 pseudomonas
821	7	1.4	445	16	Q92BW9	Q92bw9 streptomyc	894	7	1.4	499	4	Q9NRH5	Q9NRH5 homo sapien
822	7	1.4	445	16	Q9AKA2	Q9AKA2 streptomyc	895	7	1.4	499	4	Q9NRH4	Q9NRH4 homo sapien
823	7	1.4	446	16	Q9R849	Q9R849 kluyveromyc	896	7	1.4	500	4	Q9NPE0	Q9NPE0 homo sapien
824	7	1.4	446	16	Q9B1G7	Q9B1G7 rhizobium 1	897	7	1.4	500	4	Q9NPE2	Q9NPE2 homo sapien
825	7	1.4	446	16	Q9B1G7	Q9B1G7 rhizobium 1	898	7	1.4	500	4	Q9NRH8	Q9NRH8 homo sapien
826	7	1.4	447	16	Q9B1V1	Q9B1V1 argiotepe au	899	7	1.4	500	4	Q9NRH6	Q9NRH6 homo sapien
827	7	1.4	448	2	Q9RPP3	Q9RPP3 burkholderi	900	7	1.4	500	4	Q9NRH7	Q9NRH7 homo sapien
828	7	1.4	449	2	Q9S555	Q9S555 legionella	901	7	1.4	500	5	Q9NRH9	Q9NRH9 leishmania
829	7	1.4	450	16	Q9A7J6	Q9A7J6 caulobacter	902	7	1.4	500	10	Q94DZ0	Q94DZ0 oryza sativ
830	7	1.4	451	16	Q9S5B8	Q9S5B8 synechocyst	903	7	1.4	501	2	Q9VQP4	Q9VQP4 burkholderi
831	7	1.4	451	16	Q9KPS9	Q9KPS9 vibrio chol	904	7	1.4	503	5	Q21698	Q21698 caenorhabdi
832	7	1.4	452	5	Q19431	Q19431 caenorhabdi	905	7	1.4	503	16	Q84153	Q84153 chlamydia t
833	7	1.4	454	16	Q8XOV1	Q8XOV1 ralstonia s	906	7	1.4	504	5	Q95TV1	Q95TV1 homo sapien
834	7	1.4	456	16	Q9J1L9	Q9J1L9 mus musculi	907	7	1.4	504	16	Q9ZBM5	Q9ZBM5 streptomyce
835	7	1.4	457	16	Q8X673	Q8X673 escherichia	908	7	1.4	505	11	Q9ESH7	Q9ESH7 rattus norv
836	7	1.4	457	16	Q53893	Q53893 mycobacteri	909	7	1.4	507	2	Q9EXL0	Q9EXL0 streptomyce
837	7	1.4	457	16	Q8XW97	Q8XW97 oncorhynch	910	7	1.4	507	16	Q50506	Q50506 oryza sativ
838	7	1.4	461	13	Q902G3	Q9X175 vibrio chol	911	7	1.4	508	10	Q9S5Z6	Q9S5Z6 mus musculi
839	7	1.4	462	16	Q9KR75	P95Z24 mycobacteri	912	7	1.4	508	11	Q91W40	Q91W40 halobacteri
840	7	1.4	463	16	P95Z24	Q93H16 streptomyce	913	7	1.4	509	17	Q9HNR0	Q9HNR0 halobacteri
841	7	1.4	465	2	Q93H16	Q93H16 streptomyce	914	7	1.4	511	16	Q9A4C3	Q9A4C3 caulobacter
842	7	1.4	465	4	P78457	P78457 homo sapien	915	7	1.4	512	15	Q91079	Q91079 human immun
843	7	1.4	467	16	Q06Z56	Q06Z56 mycobacteri	916	7	1.4	513	2	Q24719	Q24719 comamonas t
844	7	1.4	468	16	Q8RCF0	Q8RCF0 thermosoma	917	7	1.4	513	2	Q93HP2	Q93HP2 streptomyce
845	7	1.4	469	5	Q76237	Q76237 crypanosoma	918	7	1.4	513	2	Q46334	Q46334 comamonas t
846	7	1.4	469	17	Q9YCH1	Q9YCH1 aeterypum p	919	7	1.4	515	2	Q95IF9	Q95IF9 homo sapien
847	7	1.4	470	9	Q8SCJ5	Q8SCJ5 bacterioph	920	7	1.4	515	2	Q91E79	Q91E79 streptomyce
848	7	1.4	470	10	Q8S721	Q8S721 oryza sativ	921	7	1.4	515	10	Q91E80	Q91E80 oryza sativ
849	7	1.4	470	16	Q8ZNM0	Q8ZNM0 salmonella	922	7	1.4	516	16	Q8W080	Q8W080 oryza sativ
850	7	1.4	470	16	Q8Z5F5	Q8Z5F5 salmonella	923	7	1.4	517	16	Q91L050	Q91L050 streptomyce
851	7	1.4	470	16	Q8YCS1	Q8YCS1 brucella me	924	7	1.4	518	2	Q9F8A4	Q9F8A4 treponema p
852	7	1.4	471	10	Q8SBA0	Q8SBA0 oryza sativ	925	7	1.4	519	5	Q9W131	Q9W131 drosophila
853	7	1.4	471	13	Q93314	Q93314 oncorhynch	926	7	1.4	520	2	Q9ZF72	Q9ZF72 burkholderi
854	7	1.4	472	5	Q9VF13	Q9VF13 drosophila	927	7	1.4	520	3	Q12610	Q12610 emerisella
855	7	1.4	473	5	Q9GSB0	Q9GSB0 dictyosteli	928	7	1.4	520	10	Q946Y1	Q946Y1 gossypium a
856	7	1.4	473	16	Q9RFP2	Q9RFP2 streptomyce	929	7	1.4	521	16	Q96FP6	Q96FP6 rhizobium 1
857	7	1.4	473	16	Q93GN9	Q93GN9 salmonella	930	7	1.4	521	16	Q93J76	Q93J76 streptomyce
858	7	1.4	474	5	Q20694	Q20694 caenorhabdi	931	7	1.4	522	15	Q9N4W5	Q9N4W5 caenorhabdi
859	7	1.4	475	10	Q9FMN3	Q9FMN3 streptomyce	932	7	1.4	523	16	Q9KYG3	Q9KYG3 streptomyce
860	7	1.4	475	10	Q9FMN3	Q9FMN3 oryza sativ	933	7	1.4	524	16	Q8U4T9	Q8U4T9 halobacteri
861	7	1.4	475	10	Q23233	Q23233 arabidopsis	934	7	1.4	524	16	Q91CB5	Q91CB5 bacillus su
862	7	1.4	475	12	Q55757	Q55757 chilo iride	935	7	1.4	524	16	Q9KUG5	Q9KUG5 streptomyce
863	7	1.4	476	13	Q93315	Q93315 cyprinus ca	936	7	1.4	524	16	Q9ASN3	Q9ASN3 rhizobium 1
864	7	1.4	476	16	Q8XA75	Q8XA75 escherichia	937	7	1.4	525	16	Q9E4G0	Q9E4G0 arabidopsis
865	7	1.4	476	16	Q8VKAS	Q8VKAS mycobacteri	938	7	1.4	525	16	Q9C868	Q9C868 triticum ae
866	7	1.4	477	16	Q9RW29	Q9RW29 delnoccocus	939	7	1.4	527	10	Q9AUJ7	Q9AUJ7 streptomyce
867	7	1.4	478	12	Q91TL3	Q91TL3 lupala herp	940	7	1.4	527	16	Q9KXY9	Q9KXY9 escherichia
868	7	1.4	478	13	Q93313	Q93313 sparus aura	941	7	1.4	528	4	Q9H6U3	Q9H6U3 homo sapien
869	7	1.4	479	13	Q91AM4	Q91AM4 sparus aura	942	7	1.4	528	4	Q9H3F7	Q9H3F7 homo sapien
870	7	1.4	479	16	Q9A7M6	Q9A7M6 caulobacter	943	7	1.4	528	10	Q9FNM3	Q9FNM3 arabidopsis
871	7	1.4	479	16	Q9R6X5	Q9R6X5 anabaena sp	944	7	1.4	528	10	Q93V63	Q93V63 streptomyce
872	7	1.4	479	16	Q9F313	Q9F313 streptomyce	945	7	1.4	528	16	Q9EWM0	Q9EWM0 oryza sativ
873	7	1.4	480	11	Q9QYU3	Q9QYU3 rattus rat	946	7	1.4	530	10	Q9LJ18	Q9LJ18 mus musculi
874	7	1.4	480	11	Q64630	Q64630 arabidopsis	947	7	1.4	530	16	Q88Z76	Q88Z76 streptomyce
875	7	1.4	481	10	Q64630	Q64630 rattus rat	948	7	1.4	530	16	Q9A814	Q9A814 caulobacter
876	7	1.4	482	4	Q8RAT1	Q8RAT1 homo sapien	949	7	1.4	532	17	Q91DE4	Q91DE4 aeterypum p
877	7	1.4	482	10	Q941B8	Q941B8 arabidopsis	950	7	1.4	532	17	Q91DE4	Q91DE4 streptomyce
878	7	1.4	482	10	Q941B8	Q941B8 arabidopsis	951	7	1.4	533	16	Q53903	Q53903 pseudotabie
879	7	1.4	482	10	Q949Y0	Q949Y0 arabidopsis	952	7	1.4	534	12	Q83419	Q83419 rhizobium 1
880	7	1.4	482	10	Q8RV20	Q8RV20 oryza sativ	953	7	1.4	540	2	P72288	P72288 caenorhabdi
881	7	1.4	483	5	Q9XV67	Q9XV67 caenorhabdi	954	7	1.4	541	5	Q9XTC8	Q9XTC8 streptomyce
882	7	1.4	483	16	Q8XYV6	Q8XYV6 anabaena sp	955	7	1.4	541	16	Q9A4N7	Q9A4N7 caulobacter
883	7	1.4	483	16	Q8XOP1	Q8XOP1 ralstonia s	956	7	1.4	542	10	Q8KXK7	Q8KXK7 streptomyce
884	7	1.4	486	16	Q9RKK9	Q9RKK9 streptomyce	957	7	1.4	542	16	Q9KVB5	Q9KVB5 oryza sativ
885	7	1.4	486	2	Q931X8	Q931X8 streptomyce	958	7	1.4	542	16	Q8UJ47	Q8UJ47 agrobacteri
886	7	1.4	489	5	Q9YWD5	Q9YWD5 drosophila	959	7	1.4	544	4	Q9UB19	Q9UB19 homo sapien
887	7	1.4	489	16	Q9RTW3	Q9RTW3 vibrio chol	960	7	1.4	544	5	Q46171	Q46171 nephila cia
888	7	1.4	489	16	Q98516	Q98516 rhizobium 1	961	7	1.4	545	5	Q96570	Q96570 drosophila
889	7	1.4	490	10	Q9C8H9	Q9C8H9 arabidopsis	962	7	1.4	545	5	Q27924	Q27924 drosophila
890	7	1.4	490	16	Q9F3N7	Q9F3N7 streptomyce	963	7	1.4	547	10	Q9YDWB	Q9YDWB streptomyce
891	7	1.4	493	10	Q857B7	Q857B7 oryza sativ	964	7	1.4	547	10	Q94D75	Q94D75 oryza sativ
892	7	1.4	494	2	Q9ZNM2	Q9ZNM2 aetiomonas s	965	7	1.4	547	16	Q98HBA	Q98HBA rhizobium 1

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: January 27, 2003, 16:34:03 ; Search time 31 Seconds
(without alignments)
3296.754 Million cell updates/sec

Title: US-09-889-314-2

Perfect score: 496
Sequence: 1 DTFMSISSSGPDNKNINS.....LKAYAAISGALGAHKHTNPF 496

Scoring table: OLIGO
Gapop 60.0 , Gapext 60.0

Searched: 671580 seqs, 206047115 residues

Word size: 0

Total number of hits satisfying chosen parameters: 671580

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database:

1: sp.archaea:*
2: sp.bacteria:*
3: sp.fungi:*
4: sp.human:*
5: sp.invertebrate:*
6: sp.mammal:*
7: sp.mhc:*
8: sp.organelle:*
9: sp.phage:*
10: sp.plant:*
11: sp.rodent:*
12: sp.virus:*
13: sp.vertebrate:*
14: sp.unclassified:*
15: sp.virus:*
16: sp.bacteriap:*
17: sp.archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	493	99.4	493	16	Q92797 chlamydia p
2	487	98.2	488	16	Q9J566 chlamydia p
3	11	2.2	1937	5	Q9W3L3 drosophila
4	10	2.0	128	12	Q98279 molluscum c
5	10	2.0	137	15	Q36896 human immun
6	10	2.0	487	10	Q9P655 oryza sativ
7	9	1.8	103	10	Q8S076 oryza sativ
8	9	1.8	106	3	Q9HFC7 candida alb
9	9	1.8	109	3	Q8TFM9 fusarium cu
10	9	1.8	111	3	Q9C325 podospira a
11	9	1.8	128	12	Q8V7C9 ct virus
12	9	1.8	143	4	Q76068 homo sapien
13	9	1.8	147	10	Q94F07 arabisdopsis
14	9	1.8	149	10	Q94F02 arabisdopsis
15	9	1.8	153	10	Q9LY07 arabisdopsis
16	9	1.8	153	10	Q94FR2 arabidopsis

Done
Shane
2/5/99

17	9	1.8	169	4	Q13850	Q13850 homo sapien
18	9	1.8	169	4	Q96TF6	Q96TF6 homo sapien
19	9	1.8	178	3	Q9UR55	Q9UR55 Kluyveromyc
20	9	1.8	190	4	Q96LC3	Q96LC3 homo sapien
21	9	1.8	239	2	Q9S440	Q9S440 streptomyc
22	9	1.8	243	16	Q9S2Y8	Q9S2Y8 streptomyc
23	9	1.8	285	2	Q9RCT2	Q9RCT2 streptomyc
24	9	1.8	285	2	Q9XMT2	Q9XMT2 streptomyc
25	9	1.8	307	8	Q9XMT2	Q9XMT2 ascaris suu
26	9	1.8	308	4	Q9NPY2	Q9NPY2 homo sapien
27	9	1.8	474	16	Q8UJA1	Q8UJA1 agrobacteri
28	9	1.8	477	10	Q9SEF5	Q9SEF5 arabidopsis
29	9	1.8	487	16	Q84582	Q84582 chlamydia t
30	9	1.8	491	5	Q95TX7	Q95TX7 drosophila
31	9	1.8	491	16	Q9PJG2	Q9PJG2 chlamydia m
32	9	1.8	502	4	Q96LC4	Q96LC4 homo sapien
33	9	1.8	502	4	Q9BW38	Q9BW38 homo sapien
34	9	1.8	555	4	Q9NMM0	Q9NMM0 homo sapien
35	9	1.8	555	4	Q96CT3	Q96CT3 homo sapien
36	9	1.8	555	11	Q99K82	Q99K82 mus musculu
37	9	1.8	575	12	Q91TN6	Q91TN6 tupala herp
38	9	1.8	696	5	Q9VZY2	Q9VZY2 drosophila
39	9	1.8	746	10	Q91TD0	Q91TD0 arabidopsis
40	9	1.8	748	10	Q8W0N1	Q8W0N1 oryza sativ
41	9	1.8	810	11	Q9ES29	Q9ES29 mus musculu
42	9	1.8	950	2	Q50470	Q50470 mycobacteri
43	9	1.8	1013	3	Q96OM0	Q96OM0 drosophila
44	9	1.8	1017	16	Q06586	Q06586 mycobacteri
45	9	1.8	1019	5	Q9W268	Q9W268 drosophila
46	9	1.8	1024	5	Q9W268	Q9W268 caenorhabdi
47	9	1.8	1475	5	Q9W4G4	Q9W4G4 mycobacteri
48	9	1.8	1620	16	P96285	P96285 mycobacteri
49	9	1.8	2484	5	Q9U347	Q9U347 caenorhabdi
50	9	1.8	2607	5	Q23187	Q23187 caenorhabdi
51	9	1.8	2639	5	Q76786	Q76786 anthraea p
52	9	1.8	2655	5	Q964F4	Q964F4 anthraea y
53	9	1.6	26	4	Q9UPD8	Q9UPD8 homo sapien
54	9	1.6	60	13	Q92057	Q92057 gallus gall
55	8	1.6	67	3	Q94112	Q94112 pneumocysti
56	8	1.6	90	11	Q8R412	Q8R412 mus musculu
57	8	1.6	102	16	Q8U5V9	Q8U5V9 agrobacteri
58	8	1.6	106	3	Q94018	Q94018 candida alb
59	8	1.6	109	4	Q9N864	Q9N864 homo sapien
60	8	1.6	109	11	Q9J772	Q9J772 mus musculu
61	8	1.6	110	4	Q96UQ7	Q96UQ7 rhodocorula
62	8	1.6	129	4	Q43180	Q43180 homo sapien
63	8	1.6	140	10	Q9AY45	Q9AY45 oryza sativ
64	8	1.6	143	16	Q9HX45	Q9HX45 pseudomonas
65	8	1.6	145	16	Q9RRG3	Q9RRG3 delinococcus
66	8	1.6	157	10	Q9AMS9	Q9AMS9 oryza sativ
67	8	1.6	159	11	Q9C025	Q9C025 mus musculu
68	8	1.6	161	3	Q96WP4	Q96WP4 aspergillus
69	8	1.6	164	4	Q9NS18	Q9NS18 homo sapien
70	8	1.6	165	10	Q9AXF4	Q9AXF4 avicennia m
71	8	1.6	165	10	Q94G11	Q94G11 porteresia
72	8	1.6	165	10	Q945E8	Q945E8 pennicysti
73	8	1.6	166	3	Q9UVDO	Q9UVDO pneumocysti
74	8	1.6	168	5	Q9NER1	Q9NER1 leishmania
75	8	1.6	178	5	Q45197	Q45197 caenorhabdi
76	8	1.6	178	16	Q9AAX0	Q9AAX0 caulobacter
77	8	1.6	179	10	Q9W0L8	Q9W0L8 arabidopsis
78	8	1.6	209	16	Q06630	Q06630 mycobacteri
79	8	1.6	212	10	Q9FTR1	Q9FTR1 oryza sativ
80	8	1.6	213	5	Q9U5D5	Q9U5D5 drosophila
81	8	1.6	213	5	Q18333	Q18333 drosophila
82	8	1.6	214	16	Q8UP60	Q8UP60 agrobacteri
83	8	1.6	223	16	Q9F2R7	Q9F2R7 streptomyc
84	8	1.6	225	2	Q8RNX8	Q8RNX8 propionibac
85	8	1.6	253	5	Q9B1T4	Q9B1T4 nephila sen
86	8	1.6	254	4	Q9H8P1	Q9H8P1 homo sapien
87	8	1.6	254	16	Q9XAT2	Q9XAT2 streptomyc
88	8	1.6	262	16	Q50464	Q50464 mycobacteri
89	8	1.6	273	2	Q9K188	Q9K188 bacillus ce

THIS PAGE BLANK (USPTO)

```

CC -1- SUBCELLULAR LOCATION: Extracellular.
CC -1- DOMAIN: Highly repetitive protein characterized by regions of
CC polyalanine and glycine-rich repeating units.
CC -1- SIMILARITY: BELONGS TO THE SILK FIBROIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M37137; AAA29380.1; -
DR EMBL: U03848; AAB60212.1; -
DR PIR: A36068; A36068.
KW Silk, Repeat.
FT NON_TER 1 1
FT DOMAIN 1 655
FT REPEAT 1 25
FT REPEAT 2 25
FT REPEAT 26 38
FT REPEAT 39 66
FT REPEAT 67 96
FT REPEAT 97 130
FT REPEAT 131 158
FT REPEAT 159 191
FT REPEAT 192 204
FT REPEAT 205 235
FT REPEAT 236 262
FT REPEAT 263 292
FT REPEAT 293 303
FT REPEAT 306 333
FT REPEAT 334 360
FT REPEAT 361 394
FT REPEAT 395 424
FT REPEAT 425 458
FT REPEAT 459 485
FT REPEAT 486 512
FT REPEAT 513 525
FT REPEAT 526 555
FT REPEAT 556 582
FT REPEAT 583 612
FT REPEAT 613 642
FT REPEAT 643 655
FT REPEAT 662 662
FT REPEAT 672 672
FT CONFLICT 695 747
FT CONFLICT 695 747
FT CONFLICT 695 747
SQ SEQUENCE 747 AA; 60528 MW; 850E44B0D649E012 CRC64;
      (IN REF. 1).
Query Match 1.6%; Score 8; DB 1; Length 747;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 293 GAAAGCA 300
   |||||
Db 624 GAAAGCA 631

```

```

RESULT 45
YNS4_CAEEL
ID YNS4_CAEEL STANDARD; PRT; 747 AA.
AC P34588;
DR 01-FEB-1994 (Rel. 28, Created)
DR 01-FEB-1994 (Rel. 28, Last sequence update)
DR 15-JUN-2002 (Rel. 41, Last annotation update)
DE Hypothetical protein ZC21.4 in chromosome III.
GN ZC21.4.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nemata; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxId=6239;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-Bristol NZ;
RX MEDLINE=94150718; PubMed=7906398;
RA Wilson R., Ahnscough R., Anderson K., Baynes C., Berks M.,
RA Bonfield J., Burton J., Connell M., Copey T., Cooper J., Coulson A.,
RA Craxton M., Dear S., Du Z., Durbin R., Favello A., Fraser A.,
RA Fulton L., Gardner A., Green P., Hawkins T., Hillier L., Jier M.,
RA Latreille P., Lightling J., Lloyd C., Mortimore B., O'Callaghan M.,
RA Parsons J., Percy C., Rifkin L., Roopra A., Saunders D., Showkeen R.,
RA Sims M., Smaildon N., Smith A., Smith M., Sonhammer E., Staden R.,
RA Sturton J., Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K.,
RA Waterson R., Watson A., Weinstock L., Wilkinson-Sproat J.,
RA Woldman P.;
RT "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT elegans."
RL Nature 368:32-38(1994).
CC -1- SIMILARITY: CONTAINS 1 RHO-GAP DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: L16685; AAA28171.1; ALT-INIT.
DR HSSP: Q07960; IRGP.
DR WormPep: ZC21.4; CE00534.
DR InterPro: IPR000198; RhogAP.
DR Pfam: PF00620; RhogAP; 1.
DR SMART: SM00324; RhogAP; 1.
KW SMART: hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 747 AA; 80829 MW; 48E9D61EDC4C56B0 CRC64;
      (IN REF. 1).
Query Match 1.6%; Score 8; DB 1; Length 747;
Best Local Similarity 100.0%; Pred. No. 27;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 299 GAGGAAA 306
   |||||
Db 495 GAGGAAA 502

```

Search completed: January 27, 2003, 16:35:16
Job time : 35 secs


```

DR EMBL: AF054587; AAC08584.1; -.
DR Genew: HGNC:12968; ZNF179.
DR MIM: 601237; -.
DR InterPro: IPR001841; ZnfRING.
DR Pfam: PF00097; Zf-C3HC4; 1.
DR SMART: SM00184; RING; 1.
DR PROSITE: PS00518; ZF_RING_1; FALSE_NEG.
DR PROSITE: PS50089; ZF_RING_2; 1.
DR ZINC-finger.
KW ZNF-finger.
FT ZN-FING 57 98 RING-TYPE.
SQ SEQUENCE 632 AA; 68354 MW; F8D294815E65ED02 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 632;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 287 AAGAAVGA 294
DB 584 AAGAAVGA 591

RESULT 43
CICL_RAT STANDARD; PRT; 687 AA.
AC P51802;
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Chloride channel protein CLC-K2.
GN CLCNKB.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RX MEDLINE=9431614; PubMed=8041726;
RA Kieferle S., Fong P., Bens M., Vandewalle A., Jentsch T.;
RT "Two highly homologous members of the CLC chloride channel family in
RL both rat and human kidney.";
RL Proc. Natl. Acad. Sci. U.S.A. 91:6943-6947(1994).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE-Kidney;
RX MEDLINE=94292532; PubMed=8021279;
RA Adachi S., Uchida S., Hata M., Hirose M., Marumo F., Sasaki S.,
RA Ito H.;
RT "Two isoforms of a chloride channel predominantly expressed in thick
RL ascending limb of Henle's loop and collecting ducts of rat kidney.";
RL J. Biol. Chem. 269:17677-17683(1994).
CC -I- FUNCTION: VOLTAGE-GATED CHLORIDE CHANNEL. CHLORIDE CHANNELS HAVE
CC SEVERAL FUNCTIONS INCLUDING THE REGULATION OF CELL VOLUME;
CC MEMBRANE POTENTIAL STABILIZATION, SIGNAL TRANSDUCTION AND
CC TRANSEPITHELIAL TRANSPORT. MAY BE IMPORTANT IN URINARY
CC CONCENTRATING MECHANISMS.
CC -I- SUBCELLULAR LOCATION: Integral membrane protein.
CC -I- ALTERNATIVE PRODUCTS: 2 ISOFORMS: CLC-K2L (SHOWN HERE) AND CLC-
CC K2S; ARE PRODUCED BY ALTERNATIVE SPLICING.
CC -I- TISSUE SPECIFICITY: EXPRESSED PREDOMINANTLY IN THE KIDNEY.
CC EXPRESSED IN ALL SEGMENTS OF THE NEPHRON EXAMINED, INCLUDING THE
CC S2 SEGMENT AND THE GLOMERULUS.
CC -I- SIMILARITY: BELONGS TO THE CHLORIDE CHANNEL FAMILY.
CC -I- SIMILARITY: CONTAINS 2 CBS DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----

```

```

DR EMBL: Z30663; CAA83143.1; -.
DR EMBL: D26111; BAA05106.1; -.
DR InterPro: IPR000644; CBS_domain.
DR InterPro: IPR001807; Cl-channel_volt.
DR Pfam: PF00571; CBS; 2.
DR Pfam: PF00654; voltage_CLC; 1.
DR PRINTS: PR00762; CLCHANNEL.
DR SMART: SM00116; CBS; 1.
KW Ionic channel; Ion transport; Voltage-gated channel; Transmembrane;
KW CBS domain; Repeat; Alternative splicing.
FT DOMAIN 1 49 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 50 69 2 (POTENTIAL).
FT TRANSMEM 91 114 3 (POTENTIAL).
FT TRANSMEM 138 159 4 (POTENTIAL).
FT TRANSMEM 169 188 5 (POTENTIAL).
FT TRANSMEM 200 224 6 (POTENTIAL).
FT TRANSMEM 239 257 7 (POTENTIAL).
FT TRANSMEM 282 302 8 (POTENTIAL).
FT TRANSMEM 325 348 9 (POTENTIAL).
FT TRANSMEM 398 417 10 (POTENTIAL).
FT TRANSMEM 420 438 11 (POTENTIAL).
FT TRANSMEM 466 487 12 (POTENTIAL).
FT TRANSMEM 495 514 13 (POTENTIAL).
FT DOMAIN 515 645 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 646 664 13 (POTENTIAL).
FT DOMAIN 665 687 EXTRACELLULAR (POTENTIAL).
FT DOMAIN 549 604 CBS 1.
FT DOMAIN 620 678 CBS 2.
FT CARBOHYD 364 364 N-LINKED (GLCNAC... ) (POTENTIAL).
FT VARSPLIC 77 131 MISSING (IN ISOFORM CLC-K2S).
SQ SEQUENCE 687 AA; 75217 MW; 35B0BD775397880 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 687;
Best Local Similarity 100.0%; Pred. No. 26;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 286 LAAGAAVG 293
DB 201 LAAGAAVG 208

RESULT 44
SPDL_NEPCL STANDARD; PRT; 747 AA.
AC P19837;
DT 01-FEB-1991 (Rel. 17, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Spidroin 1 (dragline silk fibroin 1) (Fragment).
OS Nepheila clavipes (Ord spider).
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Arachnida; Araneae;
OC Araneomorphae; Eteleygnae; Araneoidae; Tetragnathidae; Nepheila.
OX NCBI_TaxID=6915;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90384959; PubMed=2402494;
RA Xu M., Lewis R.V.;
RT "Structure of a protein superfiber: spider dragline silk.";
RL Proc. Natl. Acad. Sci. U.S.A. 87:7120-7124(1990).
RN [2]
RP SEQUENCE OF 653-747 FROM N.A.
RX MEDLINE=94165058; PubMed=8120021;
RA Beckwith R., Arcidiacono S.;
RT "Sequence conservation in the C-terminal region of spider silk
RL proteins (Spidroin) from Nepheila clavipes (Tetragnathidae) and
RL Araneus bicentenarius (Araneidae).";
RL J. Biol. Chem. 269:6661-6663(1994).
CC -I- FUNCTION: Spiders major ampullate silk possesses unique
CC characteristics of strength and elasticity. Fibroin consists of
CC pseudocrystalline regions of antiparallel beta-sheet interspersed
CC with elastic amorphous segments.
CC -I- SUBUNIT: MAJOR SUBUNIT, WITH SPIDROIN 2, OF THE DRAGLINE SILK.
CC -----

```


Query Match 1.6%; Score 8; DB 1; Length 512;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 354 VKAIKAI 361
 Db 57 VKAIKAI 64

RESULT 38
 LEU1_LISMO

ID LEU1_LISMO STANDARD; PRT; 512 AA.

AC O8Y5R9;

DT 15-JUN-2002 (Rel. 41, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE 2-isopropylmalate synthase (EC 4.1.3.12) (Alpha-isopropylmalate synthase) (Alpha-IPM synthetase).

GN LEUA OR LMO1987.

OS Listeria monocytogenes.

OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.

NCBI_TaxID=1639;

OX NCBI

RP SEQUENCE FROM N.A.

RC STRAIN=EGD-e / Serovar 1/2a;

RX MEDLINE=21537279; PubMed=11679669;

RA Glaser P., Frangeul L., Buchrieser C., Rusnlok C., Amend A.,

Baquerio F., Berche P., Bloecker H., Brandt P., Chakraborty T.,

Charbit A., Cheouani F., Couve E., de Daruvar A., Depoux P.,

Domann E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,

Enlian K.-D., Fsihl H., Garcia-del Portillo F., Garrido P.,

Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,

Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkut G.,

Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjati H.,

Notdustek G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,

Remmel B., Rose M., Schlueker T., Simoes N., Tierrez A.,

Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;

Comparative genomics of Listeria species.;

Science 294:849-852(2001).

CC -1- FUNCTION: Catalyzes the condensation of the acetyl group of

acetyl-CoA with 3-methyl-2-oxobutanoate (2-oxoisovalerate) to form

3-carboxy-3-hydroxy-4-methylpentanoate (2-isopropylmalate).

CC CATALYTIC ACTIVITY: 3-carboxy-3-hydroxy-4-methylpentanoate + CoA =

-1- ACETYL-CoA + 3-methyl-2-oxobutanoate + H(2O).

CC -1- PATHWAY: Leucine biosynthesis; first step.

CC -1- SUBUNIT: Homotrimer (By similarity).

CC -1- SIMILARITY: BELONGS TO THE ALPHA-IPM SYNTHETASE / HOMOCITRATE

SYNTHASE FAMILY. LEUA 1 SUBFAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC -----

CC EMBL: AL591981; CAD00065.1; -

DR Listlist; LMO01987; -

DR InterPro; IPR002034; AIPM/hcIt-synth.

DR InterPro; IPR000891; HMGL-like.

DR Pfam; PF00682; HMGL-like; 1.

DR TIGRfams; TIGR00973; LeuA_bact; 1.

DR PROSITE; PS00815; AIPM_HOMOCIT_SYNTH_1; 1.

DR PROSITE; PS00816; AIPM_HOMOCIT_SYNTH_2; 1.

KW Leucine biosynthesis; Lyase; Complete proteome.

GN SEQUENCE 512 AA; 56096 MW; 5AABID633DAB791 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 512;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 354 VKAIKAI 361

Db 57 VKAIKAI 64

RESULT 39
 Y4TO_RHISN

ID Y4TO_RHISN STANDARD; PRT; 531 AA.

AC P55669;

DT 01-NOV-1997 (Rel. 35, Created)

DT 01-NOV-1997 (Rel. 35, Last sequence update)

DE Probable peptide ABC transporter periplasmic binding protein Y4TO

precursor.

GN Y4TO.

OS Rhizobium sp. (strain NGR234).

OC Plasmid sym PNR234a.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Rhizobium.

NCBI_TaxID=394;

OX NCBI

RP SEQUENCE FROM N.A.

RX MEDLINE=97305956; PubMed=9163424;

RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,

Perret X.;

RT "Molecular basis of symbiosis between Rhizobium and legumes.";

Nature 387:394-401(1997).

CC -1- FUNCTION: PROBABLY PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT

SYSTEM Y4TOPRS FOR A PEPTIDE.

CC -1- SUBCELLULAR LOCATION: Periplasmic (Potential).

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL EXTRACELLULAR SOLUTE-BINDING

PROTEIN FAMILY 5.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

between the Swiss Institute of Bioinformatics and the EMBL outstation -

the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial

entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

or send an email to license@isb-sib.ch).

CC -----

CC EMBL: AE000098; AAB91868.1; -

DR InterPro; IPR000914; SBP_bac_5.

DR Pfam; PF00496; SBP_bac_5; 1.

DR PROSITE; PS01040; SBP_BACTERIAL_5; 1.

KW Hypothetical protein; Transports; Periplasmic; Signal; Plasmid.

FT SIGNAL 1 32 POTENTIAL.

FT CHAIN 33 531 PROBABLE PEPTIDE ABC TRANSPORTER

PERIPLASMIC BINDING PROTEIN Y4TO.

SO SEQUENCE 531 AA; 58138 MW; F2C1EB3EBB48B27 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 531;
 Best Local Similarity 100.0%; Pred. No. 20;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 264 AGLAGAA 291
 Db 11 AGLAGAA 18

RESULT 40
 COX1_PHYPO

ID COX1_PHYPO STANDARD; PRT; 594 AA.

AC C07434;

DT 01-FEB-1996 (Rel. 33, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 15-JUN-2002 (Rel. 41, Last annotation update)

DE Cytochrome c oxidase polypeptide I (EC 1.9.3.1).

GN COX1.

OS Physarum polycephalum (Slime mold).

OS Mitochondrion.

OC Eukaryota; Mycetozoa; Myxogastria; Myxogastromycetidae; Physarida;

OC Physarum.

NCBI_TaxID=5791;

```

DR PIR: S20939; S20939.
DR InterPro: IPR002937; Amino_oxidase.
DR Pfam: PF01593; Amino_oxidase; 1.
KW Carotenoid biosynthesis; Oxidoreductase; FAD; Flavoprotein; NAD.
KM Membrane; Herbicide resistance; Complete proteome.
FT NP_BIND 7 23 FAD (ADP PART) (POTENTIAL).
FT VARIANT 195 195 R -> C (CONFERS RESISTANCE TO THE
SQ SEQUENCE 472 AA; 52920 MW; 81D0896DAA28758 CRC64;
HERBICIDE NORFLURAZON).

Query Match 1.6%; Score 8; DB 1; Length 472;
Best Local Similarity 100.0%; Pred. No. 18;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 280 GAGIAGIA 287
Db 7 GAGIAGIA 14
|||||

RESULT 36
ZDS_ANASP STANDARD; PRT; 479 AA.
ID ZDS_ANASP
AC Q9R6X4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Zeta-carotene desaturase (EC 1.14.99.30) (Carotene 7,8-desaturase).
CN CRQ OR ALL2382.
OS Anabaena sp. (strain PCC 7120).
OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
OX NCBI_TaxID=103690;
[1]
RP SEQUENCE FROM N.A.
RA Mann V., Hirschberg J.;
RT "Evolution of the carotenoid biosynthesis pathway: two pathways for
RT lycopene synthesis exist in Anabaena PCC7120.";
RL Submitted (Sep-1997) to the EMBL/GenBank/DBJ databases.

[2]
RP SEQUENCE FROM N.A.
RX MEDLINE=21595285; PubMed=11795840;
RA Kaneko T., Nakamura Y., Molk C.P., Kuritz T., Sasamoto S.,
RA Matenabe A., Iriyuchi M., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kohara M., Matsumoto M., Matsuno A., Muraki A.,
RA Nakazaki N., Shimo S., Sugimoto M., Takazawa M., Yamada M.,
RA Yasuda M., Tabata S.;
RT "Complete genomic sequence of the filamentous nitrogen-fixing
RT cyanobacterium Anabaena sp. strain PCC 7120.";
RL DNA Res. 8:205-213(2001).
CC -1- FUNCTION: Catalyzes the conversion of zeta-carotene to lycopene
CC via the intermediary of neurosporene. It carries out two
CC consecutive desaturations (introduction of double bonds) at
CC positions C-7 and C-7' (By similarity).
CC -1- CATALYTIC ACTIVITY: Zeta-carotene + AH(2) + O(2) = neurosporene +
CC A + 2 H(2)O.
CC -1- CATALYTIC ACTIVITY: Neurosporene + AH(2) + O(2) = lycopene + A + 2
CC H(2)O.
CC -1- COFACTOR: NAD, NADP, OR FAD (PROBABLE).
CC -1- PATHWAY: Carotenoid biosynthesis.
CC -1- SIMILARITY: BELONGS TO THE ZETA CAROTENE DESATURASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: Y15115; CAB56041.1; -.
DR EMBL: AP003589; BAB74081.1; -.
DR InterPro: IPR002937; Amino_oxidase.
DR InterPro: IPR00205; NAD_binding.
DR Pfam: PF01593; Amino_oxidase; 1.

```

```

KW Carotenoid biosynthesis; Oxidoreductase; NAD; Flavoprotein; FAD;
KM Complete proteome.
FT CONFLICT 92 MISSING (IN REF. 1).
SQ SEQUENCE 479 AA; 53579 MW; DB14BD098965F01 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 479;
Best Local Similarity 100.0%; Pred. No. 19;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 280 GAGIAGIA 287
Db 7 GAGIAGIA 14
|||||

RESULT 37
LEU1_LISTIN STANDARD; PRT; 512 AA.
ID LEU1_LISTIN
AC Q92A28;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 2-Isopropylmalate synthase (EC 4.1.3.12) (Alpha-isopropylmalate
DE synthase) (Alpha-IPM synthetase).
GN LEUOA OR LIN2094.
OS Listeria innocua.
OC Bacteria; Firmicutes; Bacillales; Listeriaceae; Listeria.
OX NCBI_TaxID=1642;
[1]
RP SEQUENCE FROM N.A.
RP STRAIN=CLIP 11262 / Serovar 6a;
RX MEDLINE=21537279; PubMed=11679669;
RA Glaser P., Frangeul L., Buchrieser C., Rusnok C., Amend A.,
RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,
RA Chabrit A., Chetoui A., Couve E., de Darvar A., Deloux P.,
RA Dornan E., Dominguez-Bernal G., Duchaud E., Durant L., Dussurget O.,
RA Ertan K.P., Fehli H., Garcia-del Portillo F., Garrido P.,
RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,
RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapat G.,
RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,
RA Nordisk G., Novella S., de Pablos B., Perez-Diaz J.-C., Purcell R.,
RA Remmel B., Rose M., Schueter T., Simoes N., Tietze A.,
RA Vazquez-Boland J.-A., Voss H., Wehland J., Cossart P.;
RT "Comparative genomics of Listeria species.";
RL Science 294:845-852(2001).
CC -1- FUNCTION: Catalyzes the condensation of the acetyl group of
CC acetyl-CoA with 3-methyl-2-oxobutanoate (2-oxoisovalerate) to form
CC 3-carboxy-3-hydroxy-4-methylpentanoate (2-isopropylmalate).
CC -1- CATALYTIC ACTIVITY: 3-carboxy-3-hydroxy-4-methylpentanoate + CoA =
CC acetyl-CoA + 3-methyl-2-oxobutanoate + H(2)O.
CC -1- PATHWAY: Leucine biosynthesis; first step.
CC -1- SUBUNIT: Homotrimer (By similarity).
CC -1- SIMILARITY: BELONGS TO THE ALPHA-IPM SYNTHETASE / HOMOCITRATE
CC SYNTHASE FAMILY. LEUOA 1 SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL596171; CAC97324.1; -.
DR EMBL: L1N02094; -.
DR InterPro: IPR002034; AIPM/HcIt_synth.
DR InterPro: IPR000891; HMG1-like.
DR Pfam: PF00682; HMG1-like; 1.
DR TIGRfam: TIGR00973; leuA_bact; 1.
DR PROSITE: PS00815; AIPM_HOMOCIT_SYNTH_1; 1.
DR PROSITE: PS00816; AIPM_HOMOCIT_SYNTH_2; 1.
DR Leucine biosynthesis; lysase; Complete proteome.
SQ SEQUENCE 512 AA; 56149 MW; 23836636815C782 CRC64;

```

DT 15-JUN-2002 (Rel. 41, Created)
 DT 15-JUN-2002 (Rel. 41, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Molybdenum cofactor synthase protein 3 (Molybdopterin synthase
 DE sulfurylase) (MPT synthase sulfurylase).
 GN MOC53.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 OC NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Scalimayer B., Coyne K.E., Wuebbens M.M., Johnson J.L.,
 RA Rajagopalan K.V., Mendel R.R.;
 RT "The cDNA sequence of MOC53, human molybdopterin synthase
 RT sulfurylase.";
 RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21638749; PubMed=11780052;
 RA Deloukas P., Matthews L.H., Ashurst J., Burton J., Gilbert J.G.R.,
 RA Jones M., Stavrides G., Almeida J.P., Babbage A.K., Baguley C.L.,
 RA Bailey J., Barlow K.F., Bates K.N., Beard L.M., Beare D.M.,
 RA Beasley O.P., Bird C.P., Blakey S.E., Bridgeman A.M., Brown A.J.,
 RA Buck D., Burrill W.D., Butler A.P., Carder C., Carter N.P.,
 RA Chapman J.C., Clamp M., Clark G., Clark L.N., Clark S.Y., Clee C.M.,
 RA Clegg S., Cobley V.E., Collier R.E., Connor R.E., Corby N.R.,
 RA Coulson A., Coville G.J., Deadman R., Dhami P.D., Dunn M.,
 RA Ellington A.G., Frankland J.A., Fraser A., French L., Garner P.,
 RA Graffham D.V., Griffiths C., Griffiths M.N.D., Gwilliam R., Hall R.E.,
 RA Hammond S., Harley J., Heath P.D., Ho S., Holden J.L., Howden P.J.,
 RA Huckle E., Hunt A.R., Hunt S.E., Jekosch K., Johnson C.M., Johnson D.,
 RA Kay M.P., Kimberley A.M., King A., Knights A., Laird G.K., Lawlor S.,
 RA Marsh V.L., Martin S.L., McConachie L.J., McElay K., McMurray A.A.,
 RA Milne S.A., Mistry D., Moore M.J.F., Mullikin J.C., Nickerson T.,
 RA Oliver K., Parker A., Patel R., Pearce T.A.V., Peck A.I.,
 RA Phillimore B.J.C.T., Prathalingam S.R., Plumb R.W., Ramsay H.,
 RA Rice C.M., Ross M.T., Scott C.E., Sehra H.K., Showkneen R., Sims S.,
 RA Skuce C.D., Smith M.L., Soderlund C., Steward C.A., Sulston J.E.,
 RA Swann R.M., Symmore N., Taylor R., Tee L., Thomas D.W., Thorpe A.,
 RA Tracey A., Tromans A.C., Vaudin M., Wall M., Wallis J.M.,
 RA Whitehead S.L., Whitlaker P., Willey D.L., Williams L., Williams S.A.,
 RA Wilmot L., Wray P.W., Hubbard T., Durbin R.M., Bentley D.R., Beck S.,
 RA Rogers J.;
 RL "The DNA sequence and comparative analysis of human chromosome 20.";
 RL Nature 414:865-871(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=uterus;
 RA Strausberg R.;
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
 CC -1- FUNCTION: Activates MPT synthase by the ATP dependant adenylation
 CC of its C-terminal residue (By similarity).
 CC -1- PATHWAY: Molybdenum cofactor biosynthesis.
 CC -1- SIMILARITY: IN THE N-TERMINAL SECTION; BELONGS TO THE
 CC HESG/NOEF/THIF FAMILY.
 CC -1- SIMILARITY: CONTAINS 1 RHODANESE DOMAIN.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; AF102544; AAC72412.1; -
 CC EMBL; AL034553; CAB53750.1; -
 CC EMBL; BC015939; AAH15939.1; -
 CC GeneW; HGNC:15765; MOC53.
 DR InterPro; IPR000205; NAD_binding.
 DR InterPro; IPR001763; Rhodanese-like.

DR InterPro; IPR000594; Thif_domain.
 DR Pfam; PF00581; Rhodanese; 1.
 DR Pfam; PF00899; Thif; 1.
 DR SMART; SM00450; RHOD; 1.
 KW Ligase; Molybdenum cofactor biosynthesis.
 FT DOMAIN 350 460
 FT DOMAIN RHODANESE.
 SQ SEQUENCE 460 AA; 49669 MW; 29944E75513E324 CRC64;
 Query Match 1.68; Score 8; DB 1; Length 460;
 Best Local Similarity 100.0%; Pred. No. 18;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Oy 137 LQSLSAQ 144
 Db 425 LQSLSAQ 432
 RESULT 35
 CRTI_SYNY3
 ID CRTI_SYNY3 STANDARD; PRT; 472 AA.
 AC P29273;
 DT 01-DEC-1992 (Rel. 24, Created)
 DT 01-FEB-1996 (Rel. 33, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phytoene dehydrogenase (EC 1.14.99.-) (Phytoene desaturase).
 DE PDS OR CRTD OR SDR1254.
 GN Synchocystis sp. (strain PCC 6803).
 OS Bacteria; Cyanobacteria; Chroococcales; Synchocystis.
 OC NCBI_TaxID=1148;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92256820; PubMed=1581575;
 RA Martinez-Ferez I.M., Vioque A.;
 RT "Nucleotide sequence of the phytoene desaturase gene from
 RT Synchocystis sp. PCC 6803 and characterization of a new mutation
 RT which confers resistance to the herbicide norflurazon.";
 RL Plant Mol. Biol. 18:981-983(1992).
 RN [2]
 RP REVISIONS TO C-TERMINUS.
 RA Vioque A.;
 RL Submitted (NOV-1993) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=97061201; PubMed=8905231;
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
 RA Miyajima N., Hikosawa M., Sugita M., Sasamoto S., Kimura T.,
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K.,
 RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,
 RA Yamada M., Yasuda M., Tabata S.;
 RT "Sequence analysis of the genome of the unicellular cyanobacterium
 RT Synchocystis sp. strain PCC6803. II. Sequence determination of the
 RT entire genome and assignment of potential protein-coding regions.";
 RL DNA Res. 3:109-136(1996).
 CC -1- FUNCTION: THIS ENZYME CONVERTS PHYTOENE INTO ZETA-CAROTENE VIA THE
 CC INTERMEDIARY OF PHYTOFLUENE BY THE SYMMETRICAL INTRODUCTION OF TWO
 CC DOUBLE BONDS AT THE C-11 AND C-11' POSITIONS OF PHYTOENE.
 CC -1- COFACTOR: NAD, NADP, OR FAD (PROBABLE).
 CC -1- ENZYME REGULATION: INHIBITED BY THE HERBICIDE NORFLURAZON IN A
 CC NON-COMPETITIVE WAY.
 CC -1- PATHWAY: Carotenoid biosynthesis.
 CC -1- SUBCELLULAR LOCATION: Membrane-associated (Probable).
 CC -1- SIMILARITY: BELONGS TO THE PHYTOENE DEHYDROGENASE FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL; X62574; CAA44452.1; -
 CC EMBL; D90909; BAA17847.1; -

```

FT TRANSMEM 14 34 POTENTIAL.
SQ SEQUENCE 344 AA; 36421 MW; AC37A4EF919B71B5 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 344;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 AAAAGAA 301
DB 49 AAAAGAA 56

RESULT 32
HYBP_ALCEU STANDARD; PRT; 361 AA.
AC P31902;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hydrogenase nickel incorporation protein hybp.
GN HYBP.
OS Alcaligenes eutrophus (Ralstonia eutropha).
OG Plasmid megaplasmid pHG1.
OC Bacteria; Proteobacteria; beta subdivision; Ralstonia group;
CC Ralstonia.
OX NCBI_TaxID=510;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-H16 / DSM 428 / ATCC 17699;
RA MEDLINE=93356597; PubMed=8352644;
RA Derrenede J., Ellinger M., Friedrich B.;
RT "Analysis of a pleiotropic gene region involved in formation of
RT catalytically active hydrogenases in Alcaligenes eutrophus H16.";
RT Arch. Microbiol. 153:545-553(1993).
RN [2]
RP REVISIONS.
RC STRAIN-H16 / DSM 428 / ATCC 17699;
RA Ellinger T.;
RL Submitted (NOV-1994) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: COULD BE INVOLVED IN NICKEL BINDING AND ACCUMULATION.
CC -1- SIMILARITY: BELONGS TO THE HYBP/HUPM FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X70183; CAA49732.1; -
DR PIR: S29976;
DR InterPro: IPR004392; Hybp.
DR InterPro: IPR002894; Hybp_Ureg.
DR Pfam: PF01495; Hybp_Ureg_1
DR TIGRFAMs: TIGR00073; Hybp_2;
DR Metal-binding; Nickel; Plasmid.
SQ SEQUENCE 361 AA; 38483 MW; BC42976B7C5F7C4 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 361;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 300 AAGAAAT 307
DB 323 AAGAAAT 330

RESULT 33
TRMU_MYCTU STANDARD; PRT; 367 AA.
ID TRMU_MYCTU
AC 053271;
DT 30-MAY-2000 (Rel. 39, Created)

```

```

DT 30-MAY-2000 (Rel. 39, Last sequence update)
DE 16-OCT-2001 (Rel. 40, Last annotation update)
DE Probable tRNA (5-methylaminomethyl-2-thiouridylate)-methyltransferase
DE (EC 2.1.1.61).
DE TRMU OR RV3024C OR MT3108 OR MFV012.39C.
GN Mycobacterium tuberculosis.
OS Mycobacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Bacteria; Actinobacteria; Corynebacterineae; Mycobacteriaceae; Mycobacterium.
CC Actinomycetales; Corynebacterineae; Mycobacterineae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37Rv;
RC MEDLINE=96295987; PubMed=9634230;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekle A.F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skellton S., Squares S., Squares R.,
RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."
RL Nature 393:537-544(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CDC 1551 / Oshkosh;
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,
RA Bishai W.;
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and
RT laboratory strains."
RL Submitted (APR-2001) to the EMBL/Genbank/DBJ databases.
CC -1- CATALYTIC ACTIVITY: S-adenosyl-L-methionine + tRNA = S-adenosyl-L-
CC homocysteine + tRNA containing 5-methylaminomethyl-2-
CC thiouridylate.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- SIMILARITY: BELONGS TO THE TRMU FAMILY.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AL021287; CAM16109.1; -
DR EMBL: AE007129; AAK47438.1; -
DR TIGR: MT3108; -
DR Tuberculist; RV3024C; -
DR InterPro: IPR004506; Trmu.
DR InterPro: IPR004135; tRNA_Me_trans.
DR Pfam: PF03054; tRNA_Me_trans_1.
DR TIGRFAMs: TIGR00420; trmu_1.
DR Transferase; Methyltransferase; tRNA processing; Complete proteome.
SQ SEQUENCE 367 AA; 38120 MW; 043ECC2D5D690C79 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 367;
Best Local Similarity 100.0%; Pred. No. 15;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 59 ATTAGAG 66
DB 351 ATTAGAG 358

RESULT 34
MOG3_HUMAN STANDARD; PRT; 460 AA.
ID MOG3_HUMAN
AC 095396;

```

```

DE Lipopolysaccharide core biosynthesis protein rfas.
GN Rfas OR B3629.
OS Escherichia coli.
CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
CC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-K12;
RX MEDLINE=92325066; PubMed=1624461;
RA Pradel E., Parker C.T., Schmittman C.A.;
RT "Structures of the rfaB, rfaI, rfaJ, and rfaS genes of Escherichia
RT coli K-12 and their roles in assembly of the lipopolysaccharide
RT core.";
RL J. Bacteriol. 174:4736-4745(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-K12 / MG1655;
RX MEDLINE=94316500; PubMed=8041620;
RA Sofia H.J., Burland V., Daniels D.L., Plunkett G. III, Blattner F.R.;
RT "Analysis of the Escherichia coli genome. V. DNA sequence of the
RT region from 76.0 to 81.5 minutes.";
RL Nucleic Acids Res. 22:2576-2586(1994).
CC -i- PATHWAY: Lipopolysaccharide core biosynthesis.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M80599; AAA24084.1; -.
DR EMBL: U00039; AB18606.1; -.
DR EMBL: AE000440; AAC7653.1; -.
DR PIR: A42982; A42982.
DR Ecogene: EG1350; rfas.
KW Lipopolysaccharide biosynthesis; Complete proteome.
SQ SEQUENCE 311 AA; 36730 MW; C03F1BEEA6DBC26 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 311;
Best Local Similarity 100.0%; Pred. No. 13;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 67 KDKTSSTT 74
|11111111
DB 191 KDKTSSTT 198

RESULT 30
PSBO_EUGR STANDARD; PRT; 338 AA.
ID PSBO_EUGR
AC P46483;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Oxygen-evolving enhancer protein 1, chloroplast precursor (OEEL).
GN PSBO.
OS Euglena gracilis.
CC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.
OX NCBI_TaxID=3039;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 94-110.
RC STRAIN-Z;
RX MEDLINE=94154233; PubMed=8111018;
RA Yamamoto Y., Inagaki J., Mori H., Nishimura M., Takahashi S.,
RA Shigemori Y.;
RT "The presence of the precursor to the nucleus-encoded 30 kDa
RT protein of photosystem II in Euglena gracilis Z includes two
RT hydrophobic domains.";
RL Plant Mol. Biol. 24:209-215(1994).
CC -i- FUNCTION: STABILIZES THE MANGANESE CLUSTER WHICH IS THE PRIMARY

```

```

CC SITE OF WATER SPLITTING (BY SIMILARITY).
CC -i- SUBUNIT: MONOMER.
CC -i- SUBCELLULAR LOCATION: Chloroplast thylakoid membrane; associated
CC with the photosystem II complex.
CC -i- SIMILARITY: BELONGS TO THE PSBO FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D14702; BAA03529.2; ALT_INTT.
DR InterPro: IPR002628; PSI1_MSP.
DR Pfam: PF01716; MSP.1.
KW Photosynthesis; Photosystem II; Chloroplast; Transit peptide;
KW Thylakoid; Membrane; Manganese.
FT TRANSIT 1 93 CHLOROPLAST.
FT CHAIN 94 338 OXYGEN-EVOLVING ENHANCER PROTEIN 1.
SQ SEQUENCE 338 AA; 35271 MW; 1034FE4010B1A2B CRC64;

Query Match 1.6%; Score 8; DB 1; Length 338;
Best Local Similarity 100.0%; Pred. No. 14;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 288 AGAANGAA 295
|11111111
DB 41 AGAANGAA 48

RESULT 31
LIC2_BURCE STANDARD; PRT; 344 AA.
ID LIC2_BURCE
AC Q92EM5;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Lipase chaperone (Lipase foldase) (Lipase helper protein)
DE (Lipase activator protein) (Lipase modulator).
GN LIPB OR HP.
OS Burkholderia cepacia (pseudomonas cepacia).
CC Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
CC Burkholderia.
OX NCBI_TaxID=292;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-ATCC 21808;
RX MEDLINE=99124623; PubMed=9925617;
RA Ouyen D.T., Schmidt-Dannert C., Schmid R.D.;
RT "High-level formation of active pseudomonas cepacia lipase after
RT heterologous expression of the encoding gene and its modified
RT chaperone in Escherichia coli and rapid in vitro refolding.";
RL Appl. Environ. Microbiol. 65:787-794(1999).
CC -i- FUNCTION: SEEMS TO BE ACTING IN THE FOLDING OF THE EXTRACELLULAR
CC LIPASE DURING ITS PASSAGE THROUGH THE PERIPLASM.
CC -i- SUBCELLULAR LOCATION: INNER-MEMBRANE ANCHORED (BY SIMILARITY).
CC -i- SIMILARITY: BELONGS TO THE LIPASE CHAPERONE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AJ131766; CAA10510.1; -.
DR InterPro: IPR004961; Lipase_chap.
DR Pfam: PF03280; Lipase_chap.1.
KW Lipid degradation; Chaperone; Transmembrane; Periplasmic;
KW Inner membrane.

```

```

CC (VCFS/DGS). BELONGS TO THE PAIRED HOMEOBOX FAMILY.
CC -1- SIMILARITY: BELONGS TO THE PAIRED HOMEOBOX FAMILY.
CC "BICOID" SUBFAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: U96402; AAC39544.1; -.
DR HSSP: P06601; 1FJL.
DR TRANSFAC: T04038; -.
DR Genew: HGNC:4613; GSCl.
DR MIM: 601845; -.
DR InterPro: IPR001356; Homeobox.
DR Pfam: PF00046; homeobox; 1.
DR ProDom: PD000010; Homeobox; 1.
DR SMART: SM00389; HOX; 1.
DR PROSITE: PS00027; HOMEOBOX_1; 1.
DR PROSITE: PS50071; HOMEOBOX_2; 1.
DR Homeobox; DNA-binding; Nuclear protein; Polymorphism.
FT DOMAIN 2 POLY-ALA.
FT DNAS_BIND 64 70 POLY-CYS.
FT VARIANT 126 185 HOMEOBOX.
FT VARIANT 47 47 R->C.
SQ SEQUENCE 205 AA; 21545 MW; 665C33D9C454E7A7 CRC64;
Query Match 1.6%; Score 8; DB 1; Length 205;
Best Local Similarity 100.0%; Pred. No. 8.9;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 294 AAAAGGAA 301
Db 2 AAAAGGAA 9
RESULT 27
RL10_EUGGR STANDARD; PRT; 215 AA.
AC Q39724;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 60S ribosomal protein L10.
GN RPL10.
OS Euglena gracilis.
OC Eukaryota; Euglenozoa; Euglenida; Euglenales; Euglena.
OX NCBI_TaxID=3039;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=2;
RA Deruere J., Schantz M.L., Schantz R.;
RA Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
RL -1- SIMILARITY: BELONGS TO THE L10F FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X94087; CAA63831.1; -.
DR InterPro: IPR001197; Ribosomal_L10E.
DR Pfam: PF00826; Ribosomal_L10e; 1.
DR TIGRFAMs: TIGR00279; L10e; 1.
DR PROSITE: PS01257; RIBOSOMAL_L10E; 1.
DR Ribosomal protein.

```

```

SQ SEQUENCE 215 AA; 24642 MW; 4B329C0C13F37B2E CRC64;
Query Match 1.6%; Score 8; DB 1; Length 215;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 149 EAVVVAAL 156
Db 145 EAVVVAAL 152
RESULT 28
VG67_HSVB STANDARD; PRT; 272 AA.
AC P28984;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 01-FEB-1994 (Rel. 28, Last annotation update)
DE Hypothetical gene 67 protein (IR6 protein).
GN 67 OR IR6.
OS Equine herpesvirus type 1 (strain Ab4p) (EHV-1), and
OS Equine herpesvirus type 1 (strain Kentucky A) (EHV-1).
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC Alphaherpesvirinae; Varicellovirus.
OX NCBI_TaxID=31520; 10329;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AB4P.
RX MEDLINE=92295566; PubMed=1318606;
RA Telford E.A.R., Watson M.S., McBride K., Davison A.J.;
RL "The DNA sequence of equine herpesvirus-1."
RT Virology 189:304-316(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Kentucky A;
RX MEDLINE=93079867; PubMed=1333117;
RA Breiden C.A., Yalamanchili R.R., Colle C.F., III, O'Callaghan D.J.;
RT "Identification and transcriptional mapping of genes encoded at the
RT IR/US junction of equine herpesvirus type 1."
RL Virology 191:649-660(1992).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M86664; AAB02512.1; -.
DR EMBL: M86664; AAB02502.1; -.
DR EMBL: M80429; AAA46076.1; -.
DR PIR: D36802;
DR PIR: B44215; B44215.
KW Hypothetical protein.
SQ SEQUENCE 272 AA; 30103 MW; 448000315F376654 CRC64;
Query Match 1.6%; Score 8; DB 1; Length 272;
Best Local Similarity 100.0%; Pred. No. 11;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 294 AAAAGGAA 301
Db 39 AAAAGGAA 46
RESULT 29
RFAS_ECOLI STANDARD; PRT; 311 AA.
AC P27126;
DT 01-AUG-1992 (Rel. 23, Created)
DT 01-AUG-1992 (Rel. 23, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)

```



```

CC -----
CC EMBL: M7332; AAA32712.1;
CC PIR: A40814;
CC InterPro: IPR002379; ATPase_Csub.
CC InterPro: IPR000245; Vac_ATPsynth_Csub.
CC Pfam: PF00137; ATP-synt_C; 2.
CC PRINTS: PR00122; VACATPASE.
CC TIGRFAMS: TIGR01100; V_ATP_synth_C; 1.
CC HydroLase: Hydrogen ion transport; ATP synthesis; Transmembrane;
CC Multigene family.
CC KM
CC FT DOMAIN 1 10 LUMENAL (POTENTIAL).
CC FT TRANSMEM 11 33 POTENTIAL.
CC FT DOMAIN 34 55 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 56 76 POTENTIAL.
CC FT DOMAIN 77 95 LUMENAL (POTENTIAL).
CC FT TRANSMEM 96 117 POTENTIAL.
CC FT DOMAIN 118 129 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 130 155 POTENTIAL.
CC FT DOMAIN 156 165 LUMENAL (POTENTIAL).
CC FT BINDING 142 142 DICYCLOHEXYLCARBODIIMIDE (POTENTIAL).
CC FT VARIANT 78 78 T -> P (IN CLONE 12).
CC FT VARIANT 89 89 L -> V (IN CLONE 93).
CC FT VARIANT 164 164 A -> V (IN CLONE 93).
CC SQ SEQUENCE 165 AA; 16621 MW; 900A664C6C1965BB CRC64;

Query Match
Best Local Similarity 1.6%; Score 8; DB 1; Length 165;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 282 GLAGLAG 289
DB 103 GLAGLAG 110

RESULT 25
VATL_ORYSA STANDARD; PRT; 165 AA.
AC 040635;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vacuolar ATP synthase 16 kDa proteolipid subunit (EC 3.6.3.14).
GN VARP-P1.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
ON NCBI_TaxID=4530;
RX NCB1_TaxID=4530;
RP SEQUENCE FROM N.A.
RA STRAIN=cv. Guang-1u-ai No.4; TISSUE=Shoot;
RA Xiao C.;
RA Thesis (1995), Fudan University, China.
RL -I- FUNCTION: PROTON-CONDUCTING PORE FORMING SUBUNIT OF THE MEMBRANE
CC INTEGRAL V0 COMPLEX OF VACUOLAR ATPASE. V-ATPASE IS RESPONSIBLE
CC FOR ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN
CC EUKARYOTIC CELLS.
CC -I- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (in) = ADP + phosphate +
CC H(+) (out).
CC -I- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A
CC PERIPHERAL CATALYTIC V1 COMPLEX (MAIN COMPONENTS: SUBUNITS A, B,
CC C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE
CC COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN: WHICH IS PRESENT
CC AS A HEXAMER THAT FORMS THE PROTON-CONDUCTING PORE).
CC -I- SUBCELLULAR LOCATION: Integral membrane protein. Vacuolar.
CC -I- MISCELLANEOUS: THIS SUBUNIT BINDS DICYCLOHEXYLCARBODIIMIDE (DCDD)
CC WHICH INHIBITS THE ATPASE (BY SIMILARITY).
CC -I- SIMILARITY: BELONGS TO THE V-ATPASE PROTEOLIPID SUBUNIT FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

```

```

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: U27098; AAA68175.1;
CC InterPro: IPR002379; ATPase_Csub.
CC InterPro: IPR000245; Vac_ATPsynth_Csub.
CC Pfam: PF00137; ATP-synt_C; 2.
CC PRINTS: PR00122; VACATPASE.
CC TIGRFAMS: TIGR01100; V_ATP_synth_C; 1.
CC HydroLase: Hydrogen ion transport; ATP synthesis; Transmembrane.
CC KM
CC FT DOMAIN 1 10 LUMENAL (POTENTIAL).
CC FT TRANSMEM 11 33 POTENTIAL.
CC FT DOMAIN 34 55 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 56 76 POTENTIAL.
CC FT DOMAIN 77 95 LUMENAL (POTENTIAL).
CC FT TRANSMEM 96 117 POTENTIAL.
CC FT DOMAIN 118 129 CYTOPLASMIC (POTENTIAL).
CC FT TRANSMEM 130 155 POTENTIAL.
CC FT BINDING 142 142 DICYCLOHEXYLCARBODIIMIDE (POTENTIAL).
CC FT BINDING 142 142 DICYCLOHEXYLCARBODIIMIDE (POTENTIAL).
CC SQ SEQUENCE 165 AA; 16667 MW; E580168BDD0F0FD1 CRC64;

Query Match
Best Local Similarity 1.6%; Score 8; DB 1; Length 165;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 282 GLAGLAG 289
DB 103 GLAGLAG 110

RESULT 26
GSC_L_HUMAN STANDARD; PRT; 205 AA.
ID GSC_L_HUMAN
AC 015499;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Homeobox protein goosecoid-like (GSC-2).
GN GSC_L.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
ON NCBI_TaxID=9606;
RX NCB1_TaxID=9606;
RP SEQUENCE FROM N.A., AND VARIANT CYS-47.
RX MEDLINE=98110571; PubMed=9441739;
RX Funke B., St Jore B., Puech A., Strotkin H., Edelman L., Carlson C.,
RX Rafi S., Pandita R.K., Kucherlapati R., Skoultschi A., Morrow B.E.;
RT "Characterization and mutation analysis of goosecoid-like (GSC_L), a
RT homeodomain-containing gene that maps to the critical region for
RT VCFS/DS on 22q11."
RT Genomics 46:364-372(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97294411; PubMed=9150167;
RX Gotlib S., Emanuel B.S., Driscoll D.A., Sellinger B., Wang Z.,
RX Roe B., Budarf M.L.;
RT "The digeorge syndrome minimal critical region contains a goosecoid-
RT like (GSC_L) homeobox gene that is expressed early in human
RT development."
RT Am. J. Hum. Genet. 60:1194-1201(1997).
CC -I- FUNCTION: MAY HAVE A ROLE IN DEVELOPMENT. MAY REGULATE ITS OWN
CC TRANSCRIPTION. MAY BIND THE BITCOID CONSENSUS SEQUENCE TATATC.
CC -I- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -I- TISSUE SPECIFICITY: DETECTED IN ADULT TESTIS AND PITUITARY, AND IN
CC 9-10 WK FETAL TISSUE (THORAX). PROBABLY EXPRESSED IN OTHER TISSUES
CC AT LOW LEVELS.
CC -I- DEVELOPMENTAL STAGE: EXPRESSED IN EARLY HUMAN DEVELOPMENT AS WELL
CC AS IN A LIMITED NUMBER OF ADULT TISSUES.
CC -I- DISEASE: MAY BE RELATED TO SOME OF THE DEVELOPMENTAL DEFECTS
CC ASSOCIATED WITH VELOCARDIOFACIAL AND DIGEORGE SYNDROMES

```

AC P50515: 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Vacuolar ATP synthase 16 kDa proteolipid subunit (EC 3.6.3.14).
 GN VMA3 OR SPAC1B3.14.
 OS Schizosaccharomyces pombe (fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetes; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 ON NCBI_TaxID=4896;
 RX MEDLINE=92206078; PubMed=1839480;
 RA Toyama R., Goldstein D.J., Schlegel R., Dhar R.;
 RT "A genomic sequence of the Schizosaccharomyces pombe 16 kDa vacuolar
 RT H(+)-ATPase.";
 RL Yeast 7:989-991(1991).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 KC MEDLINE=21844401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
 RA Collins M., Connor R., Cronin A., Davis P., Fellwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Voiclaert G., Aert R., Robben J., Gymnopoulos B.,
 RA Woltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borzym K., Langer I., Beck A., Leinrich H., Reinhardt R., Pohl T.M.,
 RA Eger P., Zimmermann W., Medler H., Wambutt R., Purnelle B.,
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaue V., Motlier S.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallardin C., Tallada V.A., Garzon A., Rhode G.,
 RA Daga R.R., Cruzado J.L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Paulsen I., Potshkin J.,
 RA Cerretti L., Lowe T., McCombie W.R., Paulsen I., Nurse P.;
 RA Shpakovski G.V., Ussery D., Barrell B.G., Nurse P.;
 RT "The genome sequence of Schizosaccharomyces pombe.";
 RL Nature 415:871-880(2002).
 CC -1- FUNCTION: PROTON-CONDUCTING PORE FORMING SUBUNIT OF THE MEMBRANE
 CC INTEGRAL V0 COMPLEX OF VACUOLAR ATPASE. V-ATPASE IS RESPONSIBLE
 CC FOR ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN
 CC EUKARYOTIC CELLS.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
 CC H(+) (Out).
 CC -1- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A
 CC PERIPHERAL CATALYTIC V1 COMPLEX (MAIN COMPONENTS: SUBUNITS A, B,
 CC C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE
 CC COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN: WHICH IS PRESENT
 CC AS A HEXAMER THAT FORMS THE PROTON-CONDUCTING PORE).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Vacuolar.
 CC -1- MISCELLANEOUS: THIS SUBUNIT BINDS DICYLOHEXYLCARBODIIMIDE (DCCD)
 CC WHICH INHIBITS THE ATPASE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE V-ATPASE PROTEOLIPID SUBUNIT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; X5947; CAA42572.1; -

DR EMBL; Z98598; CAB1240.1; -
 DR InterPro: IPR002379; ATPase_Csub.
 DR InterPro: IPR000245; Vac_ATPsyn_Csub.
 DR Pfam: PF00137; ATP-synt_C; 2.
 DR PRINTS: PR00122; VACATPASE.
 DR TIGRFAMs: TIGR01100; V_ATP_synt_C; 1.
 KW Hydrolyase; Hydrogen ion transport; ATP synthesis; Transmembrane.
 FT DOMAIN 1
 FT TRANSMEM 10 32 LUMENAL (POTENTIAL).
 FT DOMAIN 33 54 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 55 75 POTENTIAL.
 FT DOMAIN 76 91 LUMENAL (POTENTIAL).
 FT TRANSMEM 92 113 POTENTIAL.
 FT DOMAIN 114 125 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 126 151 POTENTIAL.
 FT DOMAIN 152 161 LUMENAL (POTENTIAL).
 FT BINDING 138 138 DICYLOHEXYLCARBODIIMIDE (POTENTIAL).
 SQ SEQUENCE 161 AA; 16325 MW; DE752DFDC8CEB639 CRC64;
 Query Match 1.6%; Score 8; DB 1; Length 161;
 Best Local Similarity 100.0%; Pred. No. 7.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 282 GLAGLAG 289
 DB 99 GLAGLAG 106
 ID VATL_AVE5A STANDARD; PRT; 165 AA.
 AC P23957;
 DT 01-MAR-1992 (Rel. 21, Created)
 DT 01-MAR-1992 (Rel. 21, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Vacuolar ATP synthase 16 kDa proteolipid subunit (EC 3.6.3.14).
 GN VATP-P1.
 OS Avena sativa (Oat).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooidae;
 OC Avenae; Avena
 ON NCBI_TaxID=4498;
 RX SEQUENCE FROM N.A.
 RC STRAIN=CV. LANG; PubMed=1831453;
 RA Lai S., Watson J.C., Hansen J.N., Sze H.;
 RT "Molecular cloning and sequencing of cDNAs encoding the proteolipid
 RT subunit of the vacuolar H(+)-ATPase from a higher plant.";
 RL J. Biol. Chem. 266:16078-16084(1991).
 CC -1- FUNCTION: PROTON-CONDUCTING PORE FORMING SUBUNIT OF THE MEMBRANE
 CC INTEGRAL V0 COMPLEX OF VACUOLAR ATPASE. V-ATPASE IS RESPONSIBLE
 CC FOR ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN
 CC EUKARYOTIC CELLS.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
 CC H(+) (Out).
 CC -1- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A
 CC PERIPHERAL CATALYTIC V1 COMPLEX (MAIN COMPONENTS: SUBUNITS A, B,
 CC C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE
 CC COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN: WHICH IS PRESENT
 CC AS A HEXAMER THAT FORMS THE PROTON-CONDUCTING PORE).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Vacuolar.
 CC -1- MISCELLANEOUS: THIS SUBUNIT BINDS DICYLOHEXYLCARBODIIMIDE (DCCD)
 CC WHICH INHIBITS THE ATPASE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE V-ATPASE PROTEOLIPID SUBUNIT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----

KW Hypothetical protein; Membrane; Lipoprotein; Signal;
 KW Complete proteome.
 FT SIGNAL 1 19 POTENTIAL.
 FT CHAIN 20 132 HYPOTHETICAL LIPOPROTEIN AQ.615.
 FT LIPID 20 20 N-ACYL DIGLYCERIDE (POTENTIAL).
 FT SEQUENCE 132 AA; 13815 MM; A303B3203A67F0AD CRC64;
 SO
 Query Match 1.6%; Score 8; DB 1; Length 132;
 Best Local Similarity 100.0%; Pred. No. 6;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 289 GAAVGAAA 296
 DB 35 GAAVGAAA 42
 RESULT 21
 NRDI_CORAM STANDARD; PRT; 144 AA.
 ID NRDI_CORAM 069272;
 AC 30-MAY-2000 (Rel. 39, Created)
 DT 30-MAY-2000 (Rel. 39, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE NrdI protein.
 GN NRDI.
 OS Corynebacterium ammoniagenes (Brevibacterium ammoniagenes).
 CC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 CC Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;
 CC Corynebacterium.
 CC NCBI_TaxID=1697;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC STRAIN=ATCC 6872;
 CC MEDLINE=96136125; PubMed=9468481;
 CC Fieschi F., Torrents E., Toulkhouva L., Jordan A., Hellman U.,
 RA Barde J., Gilbert I., Karlsson M., Sjoberg B.M.;
 RT "The manganese-containing ribonucleotide reductase of Corynebacterium
 RT ammoniagenes is a class Ib enzyme."
 RL J. Biol. Chem. 273:4329-4337(1998).
 CC -1- FUNCTION: NOT KNOWN, PROBABLY INVOLVED IN RIBONUCLEOTIDE REDUCTASE
 CC FUNCTION.
 CC -1- SIMILARITY: BELONGS TO THE NRDI FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; Y09572; CAA70764.1; -;
 DR InterPro: IPR004465; NRDI.
 DR TrEMBL; TIGR00333; nrDI; 1.
 SO SEQUENCE 144 AA; 15897 MM; FEID345DB629D251 CRC64;
 Query Match 1.6%; Score 8; DB 1; Length 144;
 Best Local Similarity 100.0%; Pred. No. 6.5;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 219 KIGLEKQA 226
 DB 136 KIGLEKQA 143
 RESULT 22
 VATL_NEUCR STANDARD; PRT; 161 AA.
 ID VATL_NEUCR P31413;
 AC 01-JUL-1993 (Rel. 26, Created)
 DT 01-JUL-1993 (Rel. 26, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Vacuolar ATP synthase 16 kDa proteolipid subunit (EC 3.6.3.14).

GN VMA-3 OR 12F11.130.
 OS Neurospora crassa.
 CC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 CC Sordariales; Sordariaceae; Neurospora.
 CC NCBI_TaxID=5141;
 CC [1]
 CC SEQUENCE FROM N.A.
 CC MEDLINE=93015814; PubMed=1400281;
 RX Bowman B.J., Sista H.;
 RA "The vacuolar ATPase of Neurospora crassa."
 RT J. Bioenerg. Biomembr. 24:361-370(1992).
 RL [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=74-OR23-1A;
 RA Schulte U., Aign V., Hehseisel J., Brandt P., Fartmann B., Holland R.,
 RA Nyakatura G., Mewes H.-W., Mannhaupt G.;
 RL Submitted (DEC-2000) to the EMBL/Genbank/DBJ databases.
 CC -1- FUNCTION: PROTON-CONDUCTING PORE FORMING SUBUNIT OF THE MEMBRANE
 CC INTEGRAL V0 COMPLEX OF VACUOLAR ATPASE. V-ATPASE IS RESPONSIBLE
 CC FOR ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN
 CC EUKARYOTIC CELLS.
 CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+) (In) = ADP + phosphate +
 CC H(+) (Out).
 CC -1- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A
 CC PERIPHERAL CATALYTIC V1 COMPLEX (MAIN COMPONENTS: SUBUNITS A, B,
 CC C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE
 CC COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN, WHICH IS PRESENT
 CC AS A HEXAMER THAT FORMS THE PROTON-CONDUCTING PORE).
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Vacuolar.
 CC -1- MISCELLANEOUS: THIS SUBUNIT BINDS DICYCLOHEXYLCARBODIIMIDE (DCCD)
 CC WHICH INHIBITS THE ATPASE (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE V-ATPASE PROTEOLIPID SUBUNIT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; L07105; AAA19974.1; -;
 DR EMBL; AL451017; CAC18222.1; -;
 DR InterPro: IPR002379; ATPase_Csub.
 DR InterPro: IPR000245; Vac_ATPsynth_Csub.
 DR Pfam: PF00137; ATP-synt_C; 2.
 DR PRINTS; PRO0122; VACATPASE.
 DR TrEMBL; TIGR01100; V_ATP-synt_C; 1.
 KW Hydrolyase; Hydrogen ion transport; ATP synthesis; Transmembrane.
 FT DOMAIN 1 8 LUMENAL (POTENTIAL).
 FT TRANSMEM 9 31 LUMENAL (POTENTIAL).
 FT DOMAIN 32 53 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 54 74 LUMENAL (POTENTIAL).
 FT DOMAIN 75 91 LUMENAL (POTENTIAL).
 FT TRANSMEM 92 113 CYTOPLASMIC (POTENTIAL).
 FT DOMAIN 114 125 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 126 151 LUMENAL (POTENTIAL).
 FT DOMAIN 152 161 DICYCLOHEXYLCARBODIIMIDE (POTENTIAL).
 FT BINDING 138 138
 SO SEQUENCE 161 AA; 16329 MM; D7C54CD209373EB4 CRC64;
 Query Match 1.6%; Score 8; DB 1; Length 161;
 Best Local Similarity 100.0%; Pred. No. 7.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 282 GIAGLAAG 289
 DB 99 GIAGLAAG 106
 RESULT 23
 VATL_SCHPO STANDARD; PRT; 161 AA.
 ID VATL_SCHPO

OY 294 AAAAGGA 301
 DB 72 AAAAGGA 79

RESULT 18

RLA3_CLAHE STANDARD: PRT: 111 AA.

ID RLA3_CLAHE
 AC P42038;
 DT 01-NOV-1995 (Rel. 32, Created)
 DT 01-NOV-1995 (Rel. 32, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 60S acidic ribosomal protein p2 (Allergen Cla h 3) (Cla h III).
 GN CLAH3.
 OS Cladosporium herbarum.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
 OC Dothideomycetes et Chaetothyriomycetes Incertae sedis;
 OC Mycosphaerellaceae; mitosporic Mycosphaerellaceae; Cladosporium.
 OX NCBI_TaxID=29918;

SEQUENCE FROM N.A.
 MEDLINE=95114396; PubMed=7814877;
 RA Zhang L., Muradia G., Curran I.H., Rode H., Vijay H.M.;
 RT "A cDNA clone coding for a novel allergen, Cla h III, of Cladosporium
 herbarum identified as a ribosomal p2 protein.";
 RT J. Immunol. 154:710-717(1995).
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
 CC PROTEIN SYNTHESIS.
 CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
 CC SUBUNIT.
 CC -1- PMM: PHOSPHORYLATED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE 112P FAMILY OF RIBOSOMAL PROTEINS.
 CC -1- CAUTION: Two distinct proteins have been termed allergen
 CC Cla h 3.

 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: X77253; CAA54470.1;
 DR InterPro: IPR001813; 60S_ribosomal;
 DR Pfam: PF00428; 60S_ribosomal; 1.
 KW Ribosomal protein; Phosphorylation; Multigene family; Allergen.
 SQ SEQUENCE 111 AA; 11124 MW; 297FA13FEEDF89 CRC64;

Query Match 1.6%; Score 8; DB 1; Length 111;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGGA 300
 DB 78 GAAAGGA 85

RESULT 19

RLA3_TRYCR STANDARD: PRT: 112 AA.

ID RLA3_TRYCR
 AC P26755;
 DT 01-AUG-1992 (Rel. 23, Created)
 DT 01-AUG-1992 (Rel. 23, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE 60S acidic ribosomal protein p2-B (P2B).
 GN Trypanosoma cruzi.
 OS Eukaryota; Euzlenozoa; Kinetoplastida; Trypanosomatidae; Trypanosoma.
 OX NCBI_TaxID=5693;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=RA;

RC

RA MEDLINE=92310999; PubMed=1614880;
 RA Vazquez M.P., Schijman A.G., Panhebra A., Levin M.J.;
 RT "Nucleotide sequence of a cDNA encoding another Trypanosoma cruzi
 RT acidic ribosomal p2 type protein (TcP2B)."
 RL Nucleic Acids Res. 20:2893-2893(1992).
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
 CC PROTEIN SYNTHESIS.
 CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
 CC SUBUNIT.
 CC -1- PMM: PHOSPHORYLATED (BY SIMILARITY).
 CC -1- SIMILARITY: BELONGS TO THE 112P FAMILY OF RIBOSOMAL PROTEINS.
 CC -1- CAUTION: IT IS UNCERTAIN WHETHER MET-1 OR MET-3 IS THE INITIATOR.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----

DR EMBL: X65065; CAA46198.1;
 DR PIR: S22950; R6UT2B.
 DR InterPro: IPR001813; 60S_ribosomal;
 DR Pfam: PF00428; 60S_ribosomal; 1.
 KW Ribosomal protein; Phosphorylation.
 SQ SEQUENCE 112 AA; 10926 MW; DE84C739B4F048F CRC64;

OY 294 AAAAGGA 301
 DB 81 AAAAGGA 88

RESULT 20

Y615_AQUAE STANDARD: PRT: 132 AA.

ID Y615_AQUAE
 AC 066867;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Hypothetical lipoprotein AQ_615 precursor.
 GN AQ_615.
 OS Aquifex aeolicus.
 OC Bacteria; Aquificae; Aquificae (class); Aquificales; Aquificaceae;
 OC Aquifex
 OX NCBI_TaxID=63363;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=VF5;
 RC MEDLINE=98196666; PubMed=9537320;
 RX Decker G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjaj M., Huber R.,
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;
 RT "The complete genome of the hyperthermophilic bacterium Aquifex
 RT aeolicus.";
 RT Nature 392:353-358(1998).
 RL Nature 392:353-358(1998).
 CC -1- SUBCELLULAR LOCATION: Attached to the membrane by a lipid anchor
 CC (potential).
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@sib-sib.ch).
 CC -----
 CC EMBL: AE000698; AAC06830.1;
 DR PROSITE: PS00013; PROKAR_LIPOPROTEIN; 1.
 DR

ID RLA4_SCHPO STANDARD; PRT: 110 AA.
AC P1478;
DT 01-AUG-1990 (Rel. 15, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60S acidic ribosomal protein p2-beta (M4).
GN RPP2B OR RPA4 OR SPBC23G7.15C.
OS Schizosaccharomyces pombe (fission yeast).
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
OC Schizosaccharomycetales; Schizosaccharomycetaceae;
OC Schizosaccharomycetes.
OX NCBI_TaxID=4896;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90220620; PubMed=2325655;
RA Beltrame M., Bianchi M.E.;
RT "A gene family for acidic ribosomal proteins in Schizosaccharomycetes
pombe: two essential and two nonessential genes.";
RL Mol. Cell. Biol. 10:2341-2348(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX STRAIN=972;
MEDLINE=21848401; PubMed=11859360;
RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
RA Squeros J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,
RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
RA Oliver K., O'Neill S., Pearson D., Quail M.A., Rabinowitch E.,
RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
RA Skelton J., Simmonds M., Squares R., Squares S., Stevens K.,
RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren R., Whitehead S.,
RA Woodward J., Volkart G., Aert R., Robben J., Grymptre B.,
RA Meijmans I., Vastereels E., Rieger M., Schaefer M., Mueller-Auer S.,
RA Gabel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
RA Borzym K., Langer I., Beck A., Lebrach H., Reinhardt R., Pohl T.M.,
RA Eger P., Zimmermann W., Medler H., Mambutt R., Purnelle B.,
RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaue V., Mottier S.,
RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
RA Lucas M., Rochet M., Galliard C., Tallada V.A., Garzon A., Rhode J.,
RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Bento J.,
RA Domiguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
RA Cerutti L., Lowe T., McCombie M.R., Paulsen I., Potashkin J.,
RA Shpakovski G.V., Ussery D., Barrett B.G., Nurse P.;
RT "The genome sequence of Schizosaccharomycetes pombe.";
RL Nature 415:871-880(2002).
CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
PROTEIN SYNTHESIS.
CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
SUBUNIT.
CC -1- PM: PHOSPHORYLATED (BY SIMILARITY).
CC -1- MISCELLANEOUS: YEASTS CONTAIN 4 INDIVIDUAL SMALL RIBOSOMAL A
PROTEINS (RPA) WHICH CAN BE CLASSIFIED INTO TWO COUPLES OF SIMILAR
BUT NOT IDENTICAL SEQUENCES. EACH COUPLE IS DISTINCTLY RELATED TO
ONE OF THE TWO A PROTEINS PRESENT IN MITOCHONDRIAL ORGANISMS.
CC -1- MISCELLANEOUS: RPA3 AND RPA4 ARE ESSENTIAL FOR CELL SURVIVAL.
CC -1- WHEREAS RPA1 AND RPA2 ARE NOT.
CC -1- SIMILARITY: ALL FOUR RPA SHOW 35% IDENTITY. RPA4 IS HIGHLY RELATED
TO RPA2.
CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M33142; AAA53337.1; -

DR EMBL; AL035065; CAA22631.1; -
DR PIR; D34715; R6BY24.
DR InterPro: IPR001813; 60S_ribosomal.
DR Pfam; PF00428; 60S_ribosomal; 1.
DR Rfam; Rfam001813; 60S_ribosomal; 1.
KW Ribosomal protein; Phosphorylation; Multigene family.
SQ SEQUENCE 110 AA; 11120 MW; 477D7B04CF3652CF CRC64;
Query Match 1.6%; Score 8; DB 1; Length 110;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 296 AAGGAGAGA 303
IIIIIIII
Db 75 AAGGAGAGA 82
RESULT 17
RLA2_ARTSA STANDARD; PRT: 111 AA.
ID RLA2_ARTSA
AC P02399;
DT 21-JUL-1986 (Rel. 01, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE 60S acidic ribosomal protein p2 (E112).
OS Artemia salina (Brine shrimp).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Crustacea;
OC Branchiopoda; Anostraca; Artemiidae; Artemia.
OX NCBI_TaxID=85549;
RN [1]
RP SEQUENCE OF 1-108 FROM N.A.
RX MEDLINE=85230659; PubMed=2839187;
RA Maassen J.A., Schop E.N., Brands J.H.G.M., van Hemert F.J.,
RA Lenstra J.A., Moller W.;
RT "Molecular cloning and analysis of cDNA sequences for two ribosomal
RT proteins from Artemia. The coordinate expression of genes for
RT ribosomal proteins and elongation factor 1 during embryogenesis of
RT Artemia.";
RL Eur. J. Biochem. 149:609-616(1985).
RN [2]
RP SEQUENCE.
RX MEDLINE=80004136; PubMed=477981;
RA Amos R., Pluijms W.J.M., Moeller W.;
RT "The primary structure of ribosomal protein e12/e12-p from Artemia
RT salina 80 S ribosomes.";
RL FEBS Lett. 104:85-89(1979).
CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
PROTEIN SYNTHESIS.
CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
SUBUNIT.
CC -1- PM: PHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; X02632; CAA26479.1; -
DR PIR; A25208; R8S812.
DR InterPro: IPR001813; 60S_ribosomal.
DR Pfam; PF00428; 60S_ribosomal; 1.
DR Rfam; Rfam001813; 60S_ribosomal; 1.
KW Ribosomal protein; Phosphorylation.
FT MOD_RES 98 98 PHOSPHORYLATION (PARTIAL).
FT CONFLICT 19 19 S -> T (IN REF. 2).
FT CONFLICT 80 80 T -> A (IN REF. 2).
SQ SEQUENCE 111 AA; 11503 MW; 69C4241E1FA4E01 CRC64;
Query Match 1.6%; Score 8; DB 1; Length 111;
Best Local Similarity 100.0%; Pred. No. 5.2;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CC -1- SIMILARITY: BELONGS TO THE V-ATPASE PROTEOLIPID SUBUNIT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: X92375; CAA63119.1; -
 DR EMBL: X92374; CAA63118.1; -
 DR EMBL: M95063; AAA18550.1; -
 DR InterPro: IPR002379; ATPase_Csdb.
 DR Pfam: PF00137; ATP-synt_C; 1.
 DR TrEMBL: TIGR01100; V_atp_synt_C; 1.
 KM Hydrolyase; Hydrogen ion transport; ATP synthesis; Transmembrane.
 FT NON_TER 1 1
 FT TRANSMEM <1 20 POTENTIAL.
 FT TRANSMEM 21 39 LUMENAL (POTENTIAL).
 FT TRANSMEM 40 61 POTENTIAL.
 FT TRANSMEM 62 73 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 74 99 POTENTIAL.
 RN DOMAIN 100 109 LUMENAL (POTENTIAL).
 FT BINDING 86 86 DICYCLOHEXYLCARBODIIMIDE (BY SIMILARITY).
 FT CONFLICT 52 52 A -> P (IN REF. 1).
 FT CONFLICT 63 63 A -> G (IN REF. 1); CAA63119).
 SQ SEQUENCE 109 AA; 11043 MW; 6B7518BD7A0FEB62 CRC64;
 Query Match 1.6%; Score 8; DB 1; Length 109;
 Best Local Similarity 100.0%; Pred. No. 5.1;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 282 GIAGIAG 289
 Db 47 GIAGIAG 54
 RESULT 15
 ID RLA2_SCHPO STANDARD; PRT: 110 AA.
 AC P08094;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 60S acidic ribosomal protein P2-alpha (A2) (L40C) (L12E1).
 GN RPP2A OR RPA2 OR SPB8B7.06.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_Taxid=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88067727; PubMed=3684587;
 RA Beltz M., Bianchi M.E.;
 RT "Sequence of the cDNA for one acidic ribosomal protein of
 RT Schizosaccharomycetes pombe."
 RL Nucleic Acids Res. 15:9089-9099(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90220620; PubMed=2325655;
 RA Beltz M., Bianchi M.E.;
 RT "A gene family for acidic ribosomal proteins in Schizosaccharomycetes
 RT pombe: two essential and two nonessential genes."
 RL Mol. Cell. Biol. 10:2341-2348(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,
 RA Oliver R., O'Neill S., Pearson D., Quail M.A., Rabinowitsch E.,
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,
 RA Skellton J., Simmonds S., Squares R., Squares S., Stevens K.,
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,
 RA Woodward J., Volkart G., Aert R., Robben J., Gymnopoulos B.,
 RA Wellens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,
 RA Gabriel C., Fuchs M., Fritze C., Holzer E., Moestl D., Hilbert H.,
 RA Borgey M., Zimmermann W., Medler H., Lehnach H., Reinhardt R., Pohl T.M.,
 RA Eger P., Cadieu E., Dreano S., Gloux S., Laureau V., Motter S.,
 RA Gallieau A., Cadieu E., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,
 RA Lucas M., Rochet M., Gallier C., Tallada V.A., Garzon A., Rhode G.,
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,
 RA Cerutti L., Lowe T., McCombie M.R., Paulsen O., Potashkin J.,
 RA Snpakowski G.V., Usery D., Barrell B.G., Nurse P.;
 RA "The genome sequence of Schizosaccharomycetes pombe."
 RA Nature 415:871-880(2002).
 [4]
 RP SEQUENCE OF 1-40.
 RX MEDLINE=8403947; PubMed=6355773;
 RA Otake E., Higo K.-H., Itoh T.;
 RT "Yeast ribosomal proteins: VII. Cytoplasmic ribosomal proteins from
 RT Schizosaccharomycetes pombe."
 RL Mol. Gen. Genet. 191:519-524(1983).
 CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
 CC PROTEIN SYNTHESIS.
 CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
 CC SUBUNIT.
 CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
 CC -1- MISCELLANEOUS: YEASTS CONTAIN 4 INDIVIDUAL SMALL RIBOSOMAL A
 CC PROTEINS (RPA) WHICH CAN BE CLASSIFIED INTO TWO COUPLES OF SIMILAR
 CC BUT NOT IDENTICAL SEQUENCES. EACH COUPLE IS DISTINCTLY RELATED TO
 CC ONE OF THE TWO A PROTEINS PRESENT IN MULTICELLULAR ORGANISMS.
 CC -1- MISCELLANEOUS: RPA3 AND RPA4 ARE ESSENTIAL FOR CELL SURVIVAL,
 CC WHEREAS RPA1 AND RPA2 ARE NOT.
 CC -1- SIMILARITY: ALL FOUR RPA SHOW 35% IDENTITY. RPA4 IS HIGHLY RELATED
 CC TO RPA2.
 CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL: Y00466; CAA68528.1; -
 DR EMBL: M3138; AAA35335.1; -
 DR EMBL: AL032684; CAA21791.1; -
 DR PIR: B34715; R6BY22.
 DR InterPro: IPR001813; 60S_ribosomal.
 DR Pfam: PF00428; 60S_ribosomal.1.
 DR Ribosomal protein; Phosphorylation; Multi-gene family.
 KW Ribosomal protein; Phosphorylation; Multi-gene family.
 SQ SEQUENCE 110 AA; 11158 MW; 2B465A46E5160FE CRC64;
 Query Match 1.6%; Score 8; DB 1; Length 110;
 Best Local Similarity 100.0%; Pred. No. 5.2;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 294 AAAAGAA 301
 Db 76 AAAAGAA 83
 RESULT 16
 ID RLA4_SCHPO
 AC P08094;
 DT 01-AUG-1988 (Rel. 08, Created)
 DT 01-AUG-1988 (Rel. 08, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE 60S acidic ribosomal protein P2-alpha (A2) (L40C) (L12E1).
 GN RPP2A OR RPA2 OR SPB8B7.06.
 OS Schizosaccharomyces pombe (Fission yeast).
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;
 OC Schizosaccharomycetes.
 OX NCBI_Taxid=4896;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=88067727; PubMed=3684587;
 RA Beltz M., Bianchi M.E.;
 RT "Sequence of the cDNA for one acidic ribosomal protein of
 RT Schizosaccharomycetes pombe."
 RL Nucleic Acids Res. 15:9089-9099(1987).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=90220620; PubMed=2325655;
 RA Beltz M., Bianchi M.E.;
 RT "A gene family for acidic ribosomal proteins in Schizosaccharomycetes
 RT pombe: two essential and two nonessential genes."
 RL Mol. Cell. Biol. 10:2341-2348(1990).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=972;
 RX MEDLINE=21848401; PubMed=11859360;
 RA Wood V., Gwilliam R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,
 RA Brooks K., Brown D., Brown S., Chillingworth T., Churcher C.M.,

RT "ORK1, a potassium-selective leak channel with two pore domains
RT cloned from *Drosophila melanogaster* by expression in *Saccharomyces*
RT cerevisiae".
RL Proc. Natl. Acad. Sci. U.S.A. 93:13256-13261(1996).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN-Berkeley:
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celiker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.C., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Branton R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos D.G.,
RA Abiri J.F., Agbayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Bokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Butlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunnov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Fertiera S., Fleischmann W.,
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Maitel B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Mekulov G., Mishra N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissensbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zhang X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*."
RL Science 287:2185-2195(2000).
CC -1- FUNCTION: BACKGROUND POTASSIUM CHANNEL. RECTIFICATION IS DEPENDENT
CC ON EXTERNAL POTASSIUM CONCENTRATION. ACTS AS AN OUTWARDLY
CC RECTIFYING CHANNEL BUT AS EXTERNAL POTASSIUM LEVELS INCREASE, THIS
CC IS REVERSED.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- TISSUE SPECIFICITY: WIDESPREAD EXPRESSION IN ADULT, STRONGEST
CC EXPRESSION IN MUSCLE, BRAIN AND OVARY. ALSO PRESENT AT LOW LEVELS
CC IN LARVA AND EMBRYO.
CC -1- MISCELLANEOUS: INHIBITED BY BARIUM.
CC -1- SIMILARITY: BELONGS TO THE TWO PORE DOMAIN FAMILY OF POTASSIUM
CC CHANNELS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL: U55321; AAC69250.1; -
CC EMBL: AE003484; AAF47972.1; -
CC FLYBase: FBgn0017561; ORK1.
CC InterPro: IPR003280; K-channel_2pore.
CC InterPro: IPR001622; K-channel_pore.
CC InterPro: IPR000636; M-channel_nlg.
CC Pfam: PF00520; Ion_trans; 1.

DR PRINTS: PRO1333; 2PORECHANNEL.
KW Ionic channel; Transmembrane; Ion transport; Potassium transport;
KW Glycoprotein.
FT DOMAIN 1 6 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 7 27 POTENTIAL.
FT DOMAIN 95 111 PORE-FORMING 1 (POTENTIAL).
FT TRANSSEM 120 140 POTENTIAL.
FT DOMAIN 141 170 CYTOPLASMIC (POTENTIAL).
FT TRANSSEM 171 191 POTENTIAL.
FT DOMAIN 208 224 PORE-FORMING 2 (POTENTIAL).
FT TRANSSEM 244 264 POTENTIAL.
FT DOMAIN 265 1001 CYTOPLASMIC (POTENTIAL).
FT CARBOHYD 58 58 N-LINKED (GLCNAC...) (POTENTIAL).
SQ SEQUENCE 1001 AA; 109289 MW; 09AE1A369072E07 CRC64;
Query Match 1.88; Score 9; DB 1; Length 1001;
Best Local Similarity 100.0%; Pred. No. 4.2;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy 294 AAAGAGAG 302
Db 764 AAAGAGAG 772
RESULT 14
VATL_MAIZE STANDARD: PRT: 109 AA.
AC 041773; Q41774; Q08074;
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Vacuolar ATP synthase 16 kDa proteolipid subunit (EC 3.6.3.14) (V-
DE ATPase 16 kDa proteolipid subunit) (Fragment).
OS Zea mays (Maize).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade;
OC Panicoideae; Andropogoneae; Zea.
OX NCBI_TaxID=4577;
RN [1]
RP SEQUENCE OF 1-76 FROM N.A.
RC STRAIN=cv. Lixis; TISSUE=coleoptile;
RX MEDLINE=96197807; PubMed=8617373;
RA Viereck R., Kirsch M., Loew R., Rausch T.;
RT "Down-regulation of plant V-type H⁺-ATPase genes after light-induced
RT inhibition of growth."
RL FEBS Lett. 364:285-288(1996).
RN [2]
RP SEQUENCE OF 52-109 FROM N.A.
RC STRAIN=cv. B73; TISSUE=leaf;
RX MEDLINE=94105294; PubMed=8278499;
RA Keith C.S., Hoang D.O., Barrett B.M., Felgelman B., Nelson M.C.,
RA "Partial sequence analysis of 130 randomly selected maize cDNA
RT clones."
RL Plant Physiol. 101:329-332(1993).
CC -1- FUNCTION: PROTON-CONDUCTING PORE FORMING SUBUNIT OF THE MEMBRANE
CC INTEGRAL V0 COMPLEX OF VACUOLAR ATPASE. V-ATPASE IS RESPONSIBLE
CC FOR ACIDIFYING A VARIETY OF INTRACELLULAR COMPARTMENTS IN
CC EUKARYOTIC CELLS.
CC -1- CATALYTIC ACTIVITY: ATP + H(2)O + H(+)(In) = ADP + phosphate +
CC H(+)(Out).
CC -1- SUBUNIT: V-ATPASE IS AN HETEROMULTIMERIC ENZYME COMPOSED OF A
CC PERIPHERAL CATALYTIC VI COMPLEX (MAIN COMPONENTS: SUBUNITS A, B,
CC C, D, E, AND F) ATTACHED TO AN INTEGRAL MEMBRANE V0 PROTON PORE
CC COMPLEX (MAIN COMPONENT: THE PROTEOLIPID PROTEIN, WHICH IS PRESENT
CC AS A HEXAMER THAT FORMS THE PROTON-CONDUCTING PORE).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Vacuolar.
CC -1- TISSUE SPECIFICITY: HIGH EXPRESSION IN THE MESOCOTYL TIP OF
CC ETIOLATED SEEDLINGS COMPARED TO THE BASE.
CC -1- DEVELOPMENTAL STAGE: EXPRESSION IS STRONGLY LINKED TO EXTENSION
CC GROWTH.
CC -1- MISCELLANEOUS: THIS SUBUNIT BINDS DICYLOHEXYLCARBODIIMIDE (DCCD)
CC WHICH INHIBITS THE ATPASE (BY SIMILARITY).

KW Capsid assembly.
 SQ SEQUENCE 782 AA; 87392 MW; 4DE40314E5BAB9B CRC64;
 Query Match 1.8%; Score 9; DB 1; Length 782;
 Best Local Similarity 100.0%; Pred. No. 3.4;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 293 GAAAGGAA 301
 |||||
 DB 436 GAAAGGAA 444

RESULT 11
 ELS_MOUSE STANDARD; PRT; 860 AA.
 AC P54320;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 01-OCT-1996 (Rel. 34, Last annotation update)
 DE Elastin precursor (tropoelastin).
 GN ELN.
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-BALB/c; TISSUE=Lung;
 RX MEDLINE=95130069; PubMed=7829060;
 RA Wyder K.S., Sechler J.L., Boyd C.D., Passmore H.C.:
 "Use of an intron polymorphism to localize the tropoelastin gene to
 mouse chromosome 5 in a region of linkage conservation with human
 chromosome 7".
 RT Genomics 23:125-131(1994).
 RL
 CC -1- FUNCTION: MAJOR STRUCTURAL PROTEIN OF TISSUES SUCH AS AORTA AND
 CC NUCAL LIGAMENT, WHICH MUST EXPAND RAPIDLY AND RECOVER COMPLETELY.
 CC -1- SUBUNIT: THE POLYMERIC ELASTIN CHAINS ARE CROSS-LINKED TOGETHER
 CC INTO AN EXTENSIBLE 3D NETWORK.
 CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX OF ELASTIC FIBERS.
 CC -1- PTM: THE CROSSLINKS ARE MADE OF DEAMINATED LYS.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U08210; AAA80155.1; -
 CC DR MGD; MGI:95317; Eln.
 CC DR InterPro; IPR003979; tropoelastin.
 CC DR PRINTS; PR01500; TROPOELASTIN.
 CC DR Structural Protein; Repeat; Signal; Connective tissue.
 CC FT SIGNAL 1 27
 CC FT CHAIN 28 860 ELASTIN.
 CC FT SEQUENCE 860 AA; 71955 MW; 0C0BE5AAE1EDD7F1 CRC64;
 SQ

Query Match 1.8%; Score 9; DB 1; Length 860;
 Best Local Similarity 100.0%; Pred. No. 3.7;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 298 GAAAGGAA 306
 |||||
 DB 113 GAAAGGAA 121

RESULT 12
 SOXA_CORSI STANDARD; PRT; 967 AA.
 ID SOXA_CORSI
 AC 046337;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Sarcosine oxidase alpha subunit (EC 1.5.3.1) (Sarcosine oxidase
 DE subunit).
 DE SOXA.
 OS Corynebacterium sp. (strain P-1).
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;
 OC Corynebacterium.
 OX NCBI_TaxID=69006;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN-P-1;
 RC MEDLINE=95355441; PubMed=7543100;
 RA Chlumsky L.J., Zhang L., Jorns M.S.:
 "Sequence analysis of sarcosine oxidase and nearby genes reveals
 RT homologues with key enzymes of folate one-carbon metabolism.";
 RL J. Biol. Chem. 270:18252-18259(1995)
 CC -1- FUNCTION: CATALYZES THE OXIDATIVE DEMETHYLATION OF SARCOSE TO
 CC YIELD GLYCINE, HYDROGEN PEROXIDE AND 5,10-
 CC METHYLENETERAHYDROFOLATE.
 CC -1- CATALYTIC ACTIVITY: Sarcosine + H(2)O + O(2) = glycine +
 CC formaldehyde + H(2)O(2).
 CC -1- COFACTOR: FAD.
 CC -1- SUBUNIT: HETEROTETRAMER (ALPHA 100 kDa, BETA 42 kDa, GAMMA 20 kDa,
 CC AND DELTA 6 kDa).
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
 CC -1- SIMILARITY: BELONGS TO THE GCVT FAMILY.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; U23955; AAC43461.1; -
 CC DR InterPro; IPR002536; GCV_T.
 CC DR InterPro; IPR000205; NAD-binding.
 CC DR Pfam; PF01571; GCV_T; 1.
 CC DR ProDom; PD000139; FAD-pyr_redox; 1.
 CC KW Oxidoreductase; Flavoprotein; FAD.
 CC SEQUENCE 967 AA; 102767 MW; C5F8A61F021134E9 CRC64;
 SQ

Query Match 1.8%; Score 9; DB 1; Length 967;
 Best Local Similarity 100.0%; Pred. No. 4.1;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 289 GAAAGGAA 297
 |||||
 DB 434 GAAAGGAA 442

RESULT 13
 ORK1_DROME STANDARD; PRT; 1001 AA.
 ID ORK1_DROME
 AC Q94526;
 DT 16-OCT-2001 (Rel. 40, Created)
 DT 16-OCT-2001 (Rel. 40, Last sequence update)
 DT 16-OCT-2001 (Rel. 40, Last annotation update)
 DE Open rectifier potassium channel protein 1 (Two pore domain potassium
 DE channel ork1).
 DE ORK1 OR CG1615.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
 OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
 OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Larva;
 RX MEDLINE=97075152; PubMed=8917578;
 RA Goldstein S.A.N., Price L.A., Rosenthal D.N., Pausch M.H.;

Db 15 GAGLAGLAA 23

RESULT 9

ELS_BOVIN STANDARD; PRT; 747 AA.

ID ELS_BOVIN STANDARD; PRT; 747 AA.

AC P04985; P04986; P04987; Q29421;

DT 13-AUG-1987 (Rel. 05, Created)

DT 13-AUG-1987 (Rel. 05, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Elastin precursor (Tropoelastin).

GN ELN.

OS Bos taurus (Bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea; Bovidae; Bovinae; Bos.

OC NCBI_TaxID=9913;

OX NCBI_TaxID=9913;

RN [1]

RX SEQUENCE FROM N.A.

RA MEDLINE=87194772; PubMed=3032943;

RA Raju K., Anwar R.A.;

RT "Primary structures of bovine elastin a, b, and c deduced from the sequences of cDNA clones."

RL J. Biol. Chem. 262:5755-5762(1987).

RN [2]

RX SEQUENCE OF 1-27 FROM N.A.

RA TISSUE=Nuchal ligament;

RA MEDLINE=89274159; PubMed=2543440;

RA Yeh H., Anderson N., Ornstein-Goldstein N., Bashir M.M., Mecham R., Rosenbloom J.C., Abrams W.R., Indik Z., Yoon K., Parks W., Mecham R., Rosenbloom J.;

RT "Structure of the bovine elastin gene and SI nuclease analysis of alternative splicing of elastin mRNA in the bovine nuchal ligament."

RL Biochemistry 28:2365-2370(1989).

RN [3]

RX SEQUENCE OF 1-27 FROM N.A.

RA MEDLINE=91234332; PubMed=2031719;

RA Manohar A., Shi W., Anwar R.A.;

RT "Partial characterization of bovine elastin gene; comparison with the gene for human elastin."

RL Biochem. Cell Biol. 69:185-192(1991).

CC -1- FUNCTION: MAJOR STRUCTURAL PROTEIN OF TISSUES SUCH AS AORTA AND NUCAL LIGAMENT, WHICH MUST EXPAND RAPIDLY AND RECOVER COMPLETELY.

CC -1- SUBUNIT: THE POLYMERIC ELASTIN CHAINS ARE CROSS-LINKED TOGETHER INTO AN EXTENSIBLE 3D NETWORK.

CC -1- SUBCELLULAR LOCATION: EXTRACELLULAR MATRIX OF ELASTIC FIBERS.

CC -1- ALTERNATIVE PRODUCTS: 3 ISOFORMS; A (SHOWN HERE), B AND C; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- PTM: THE CROSSLINKS ARE MADE OF DEAMINATED LYS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL: J02717; AAA30503.1; -

DR EMBL: K03505; AAA30505.1; -

DR EMBL: K03506; AAA30506.1; -

DR EMBL: J02855; AAA30776.1; -

DR EMBL: M58652; AAA03519.2; -

DR PIR: A26728; A26728.

DR PIR: B26728; B26728.

DR PIR: C26728; C26728.

DR InterPro: IPR003979; Tropoelastin.

DR PRINTS: PR01500; TROPELASTIN.

KW Structural protein; Connective tissue; Repeat; Signal;

KW Alternative splicing;

FT SIGNAL 1 26

FT CHAIN 27 747 ELASTIN.

FT MOD_RES 105 105 OXIDATIVE DEAMINATION.

FT MOD_RES 109 109 OXIDATIVE DEAMINATION.

FT MOD_RES 252 252 OXIDATIVE DEAMINATION.

FT MOD_RES 271 271 OXIDATIVE DEAMINATION.

FT MOD_RES 275 275 OXIDATIVE DEAMINATION.

FT MOD_RES 324 324 OXIDATIVE DEAMINATION.

FT MOD_RES 327 327 OXIDATIVE DEAMINATION.

FT MOD_RES 400 400 OXIDATIVE DEAMINATION.

FT MOD_RES 404 404 OXIDATIVE DEAMINATION.

FT MOD_RES 407 407 OXIDATIVE DEAMINATION.

FT MOD_RES 448 448 OXIDATIVE DEAMINATION.

FT MOD_RES 489 489 OXIDATIVE DEAMINATION.

FT MOD_RES 493 493 OXIDATIVE DEAMINATION.

FT MOD_RES 544 544 OXIDATIVE DEAMINATION.

FT MOD_RES 548 548 OXIDATIVE DEAMINATION.

FT MOD_RES 552 552 OXIDATIVE DEAMINATION.

FT MOD_RES 606 606 OXIDATIVE DEAMINATION.

FT MOD_RES 609 609 OXIDATIVE DEAMINATION.

FT MOD_RES 645 645 OXIDATIVE DEAMINATION.

FT MOD_RES 649 649 OXIDATIVE DEAMINATION.

FT MOD_RES 685 685 OXIDATIVE DEAMINATION.

FT MOD_RES 688 688 OXIDATIVE DEAMINATION.

FT MOD_RES 226 239 MISSING (IN ELASTIN B).

FT VARSPLIC 226 259 MISSING (IN ELASTIN C).

FT CONFLICT 1 3 MRS -> MAG (IN REF. 2 AND 3).

FT CONFLICT 12 12 E -> G (IN REF. 2 AND 3).

FT SEQUENCE 747 AA; 64229 MW; 633C03E411643D83 CRC64;

Query Match 1.8%; Score 9; DB 1; Length 747;

Best Local Similarity 100.0%; Pred. No. 3.2;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 298 GGAAGAAA 306

DB 95 GGAAGAAA 103

RESULT 10

ID PRTP_HSVT2 STANDARD; PRT; 782 AA.

AC O9MR16;

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Probable processing and transport protein.

GN PRTP.

OS Herpesvirus tupaia (Strain 2) (THV-2).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae; Betaherpesvirinae.

OC NCBI_TaxID=132678;

OX NCBI_TaxID=132678;

RN [1]

RX SEQUENCE FROM N.A.

RA MEDLINE=99319892; PubMed=10392721;

RA Bahr U., Springfeld C., Tidona C.A., Darai G.;

RT "Structural organization of a conserved gene cluster of Tupaia herpesvirus encoding the DNA polymerase, glycoprotein B, a probable processing and transport protein, and the major DNA binding protein."

RL Virus Res. 60:123-136(1999).

CC -1- FUNCTION: THIS PROTEIN MAY AFFECT TRANSLLOCATION OF THE VIRUS GLYCOPROTEINS TO MEMBRANES. IT IS INVOLVED IN CAPSID MATURATION (BY SIMILARITY).

CC -----

CC -1- SIMILARITY: BELONGS TO THE HERPESVIRUSES PRTP FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

CC -----

CC EMBL: AF084543; AADA2934.1; -

DR InterPro: IPR000501; Proc-transport.

DR Pfam: PF01366; PRTP; 1.

TRANSMEMBRANE SEGMENTS M1, M2, AND M4 FROM SUBUNITS OF THE GABA(A)/BENZODIAZEPINE RECEPTOR FAMILY, RESPECTIVELY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: M36035; AAA03652.1; -
 DR EMBL: U12421; AAA83252.1; -
 DR EMBL: Z82214; CAB55884.1; -
 DR EMBL: BC001110; AAH01110.1; -
 DR PIR: S14257; S14257.
 DR Genew: HGNC:1138; BZRP.
 DR MIM: 109610; -
 DR InterPro: IPR004307; Tspo_MBR.
 DR Pfam: PF03073; Tspo_MBR; 1.
 DR Mitochondrion: Receptor; Transmembrane; Polymorphism.

FT TRANSMEM 5 26 TM1 (POTENTIAL).
 FT TRANSMEM 47 67 TM2 (POTENTIAL).
 FT TRANSMEM 80 100 TM3 (POTENTIAL).
 FT TRANSMEM 106 126 TM4 (POTENTIAL).
 FT TRANSMEM 135 155 TM5 (POTENTIAL).
 FT TRANSMEM 147 147 A -> T.
 FT VARIANT /FTid=VAR_013617.
 FT VARIANT 162 162 H -> R.
 FT VARIANT /FTid=VAR_013618.
 FT VARIANT 162 162 IAD741BF99AB92CD CRC64;

Query Match 1.8%; Score 9; DB 1; Length 169;
 Best Local Similarity 100.0%; Pred. No. 0.89;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 302 GAAATTTVA 310
 117 GAAATTTVA 125

RESULT 7
 CRTL_STRGR STANDARD; PRT; 507 AA.
 AC P54981: P72447;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phytoene dehydrogenase (EC 1.14.99.-) (Phytoene desaturase).
 GN CRTL OR CRTE.
 OS Streptomyces griseus.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OC NCBI_TaxID=1911;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=JA3933;
 RC MEDLINE=97074881; PubMed=8917308;
 RA Schumann G., Nurnberger H., Sandmann G., Kruegel H.J.;
 RT "Activation and analysis of cryptic crt genes for carotenoid
 biosynthesis from Streptomyces griseus.";
 RT Mol. Gen. Genet. 252:656-666(1996).
 RL -1- FUNCTION: THIS ENZYME CONVERTS PHYTOENE INTO ZETA-CAROTENE VIA THE
 CC INTERMEDIARY OF PHYTOFLUENE BY THE SYMMETRICAL INTRODUCTION OF TWO
 CC DOUBLE BONDS AT THE C-11 AND C-11' POSITIONS OF PHYTOENE.
 CC -1- COFACTOR: FAD (PROBABLE).
 CC -1- PATHWAY: Carotenoid biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE PHYTOENE DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: L37405; AAA91950.1; -
 DR EMBL: X95396; CAA64850.1; -
 DR InterPro: IPR002937; Amino_oxidase.
 DR InterPro: IPR00171; Bac_phytoene_dh.
 DR InterPro: IPR000205; NAD_binding.
 DR Pfam: PF01593; Amino_oxidase; 1.
 DR PROSITE: PS00982; PHYTOENE_DH; 1.
 DR Carotenoid biosynthesis; Oxidoreductase; FAD; Flavoprotein; NAD.
 DR NP_BIND FAD (ADP PART) (POTENTIAL).
 FT NP_BIND 12 45
 FT SEQUENCE 507 AA; 54509 MW; FB97F7FE696B2AC CRC64;

Query Match 1.8%; Score 9; DB 1; Length 507;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 280 GAGLAGLAA 288
 15 GAGLAGLAA 23

RESULT 8
 CRTL_STRGR STANDARD; PRT; 508 AA.
 AC P54971;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 15-JUN-2002 (Rel. 41, Last annotation update)
 DE Phytoene dehydrogenase (EC 1.14.99.-) (Phytoene desaturase).
 GN CRTL.
 OS Streptomyces setonii.
 OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.
 OC NCBI_TaxID=38315;
 OX [1]
 RN SEQUENCE FROM N.A.
 RP STRAIN=ISP 5395;
 RA Hoshii K.;
 RT Submitted (JUL-1995) to the EMBL/Genbank/DBJ databases.
 RL -1- FUNCTION: THIS ENZYME CONVERTS PHYTOENE INTO ZETA-CAROTENE VIA THE
 CC INTERMEDIARY OF PHYTOFLUENE BY THE SYMMETRICAL INTRODUCTION OF TWO
 CC DOUBLE BONDS AT THE C-11 AND C-11' POSITIONS OF PHYTOENE.
 CC -1- COFACTOR: FAD (PROBABLE).
 CC -1- PATHWAY: Carotenoid biosynthesis.
 CC -1- SIMILARITY: BELONGS TO THE PHYTOENE DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).

DR EMBL: D55723; BAA09537.1; -
 DR InterPro: IPR002937; Amino_oxidase.
 DR InterPro: IPR00171; Bac_phytoene_dh.
 DR InterPro: IPR000205; NAD_binding.
 DR Pfam: PF01593; Amino_oxidase; 1.
 DR PROSITE: PS00982; PHYTOENE_DH; 1.
 DR Carotenoid biosynthesis; Oxidoreductase; FAD; Flavoprotein; NAD.
 DR NP_BIND FAD (ADP PART) (POTENTIAL).
 FT NP_BIND 12 45
 FT SEQUENCE 508 AA; 54610 MW; 4B6DEFC076D51CB5 CRC64;

Query Match 1.8%; Score 9; DB 1; Length 508;
 Best Local Similarity 100.0%; Pred. No. 2.3;
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB 280 GAGLAGLAA 288
 117 GAGLAGLAA 23

```

RESULT 5
RLA4 CLAHE          STANDARD:      PRT:      111 AA.
ID P42039;
DT 01-NOV-1995 (Rel. 32, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 60S acidic ribosomal protein p2 (Minor allergen Cla h 4) (Cla h IV).
GN CLA4..
OS Cladosporium herbarum.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina;
OC Didymomyces et Chaetothryiomycetes incertae sedis;
OC Mycosphaerellaceae; mitosporic Mycosphaerellaceae; Cladosporium.
OX NCBI_TaxID=29918;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=280202-Berlin;
RX MEDLINE=95206305; PubMed=7898496;
RA Achatz G., Oberkofler H., Lechenauer E., Simon B., Unger A.,
RA Kandler D., Ehner C., Prillinger H., Kraft D., Breitenbach M.;
RT "Molecular cloning of major and minor allergens of Alternaria
RT alternata and Cladosporium herbarum.";
RL Mol. Immunol. 32:213-227(1995).
RN [2]
RP REVISIONS TO 13; 38-41 AND 93.
RA Simon-Nobbe B.;
RL Submitted (Aug-1999) to the EMBL/GenBank/DDAJ databases.
CC -1- FUNCTION: PLAYS AN IMPORTANT ROLE IN THE ELONGATION STEP OF
CC PROTEIN SYNTHESIS.
CC -1- SUBUNIT: P1 AND P2 EXIST AS DIMERS AT THE LARGE RIBOSOMAL
CC SUBUNIT.
CC -1- PTM: PHOSPHORYLATED (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC -1- CAUTION: Two distinct proteins have been termed allergen
CC Cla h 4.
CC -----
CC CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sb.ch).
CC -----
CC DR EMBL: X78223; CAA55067.2; -
CC DR InterPro: IPR001813; 60S_rribosomal.
CC DR Pfam: PF00428; 60S_rribosomal; 1.
CC KW Ribosomal protein; Phosphorylation; Multigene family; Allergen.
CC SQ SEQUENCE 111 AA; 11105 MW; C7B65CGAD97876A CRC64;

Query Match      1.8%; Score 9; DB 1; Length 111;
Best Local Similarity 100.0%; Pred No. 0.61;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 293 GAAAGCAA 301
   |||||
Db 77 GAAAGCAA 85

RESULT 6
PKBS_HUMAN          STANDARD:      PRT:      169 AA.
ID P30516; Q96TF6;
AC 01-APR-1993 (Rel. 25, Created)
DT 01-FEB-1994 (Rel. 28, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE Peripheral-type benzodiazepine receptor (PBR) (PKBS) (Mitochondrial
DE benzodiazepine receptor).
GN BZRP OR MBR.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

```

```

OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91146565; PubMed=1847678;
RA Rioud J., Mattei M.-G., Kagnad M., Dumont X., Guillemot J.C.,
RA le Fur G., Caput D., Ferrara P.;
RT "Molecular cloning and chromosomal localization of a human
RT peripheral-type benzodiazepine receptor.";
RL Eur. J. Biochem. 195:305-311(1991).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=95237610; PubMed=7721091;
RA Yakovlev A.G., Ruffo M., Jurka J., Krueger K.E.;
RT "Comparison of repetitive elements in the third intron of human and
RT rodent mitochondrial benzodiazepine receptor-encoding genes.";
RL Gene 155:201-205(1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057165; PubMed=10591208;
RA Dunham I., Hunt A.R., Collins J.E., Bruskiewicz R., Beare D.M.,
RA Clump M., Smink L.J., Ainscough R., Almeida J.P., Babbage A.K.,
RA Baguley C., Bailey J., Barlow K.F., Bates K.N., Beasley O.P.,
RA Bird C.P., Blakey S.E., Bridgeman A.M., Buck D., Burgess J.,
RA Burrill W.D., Burton J., Carder C., Carter N.P., Chen Y., Clark G.,
RA Clegg S.M., Cobley V.E., Cole C.G., Collier R.E., Connor R.,
RA Conroy D., Corby N.R., Coville G.J., Cox A.V., Davis J., Dawson E.,
RA Dhani P.D., Dockree C., Dodsworth S.J., Durbin R.M., Ellington A.G.,
RA Evans K.L., Fey J.M., Fleming K., French L., Garner A.A.,
RA Gilbert J.G.R., Goward M.E., Graffham D.V., Griffiths M.N.D., Hall C.,
RA Hall R.E., Hall-Tamlyn G., Heathcote R.W., Ho S., Holmes S.,
RA Hunt S.E., Jones M.C., Kershaw J., Kimberley A.M., King A.,
RA Laird G.K., Langford C.F., Leversha M.A., Lloyd C., Lloyd D.M.,
RA Marlyn I.D., Mashregh-Mohammadi M., Matthews L.H., Mccann O.T.,
RA Mcclay J., McLaren S., McMurray A.A., Milne S.A., Mortimore B.J.,
RA Odell C.N., Pavitt R., Pearce A.V., Pearson D., Phillips M.L.,
RA Phillips S.H., Plumb R.W., Ramsay H., Ramsey Y., Rogers L., Ross M.T.,
RA Scott C.E., Sehra H.K., Skuce C.D., Smalley S., Smith M.L.,
RA Soderlund C., Spragon L., Steward C.A., Sultson J.E., Swann R.M.,
RA Vaudin M., Wall M., Wallis J.M., Whiteley M.N., Willey D.L.,
RA Williams L., Williams S.A., Williamson H., Wilmer T.E., Wilming L.,
RA Wright C.L., Hubbard T., Bentley D.R., Beck S., Rogers J., Shimizu N.,
RA Winstona S., Kawasaki K., Sasaki T., Asakawa S., Kudoh J.,
RA Shintani A., Shibuya K., Yoshizaki Y., Aoki N., Mitsuyma S.,
RA Roe B.A., Chen F., Chu L., Crabtree J., Deschamps S., Do A., Do T.,
RA Dorman A., Fang F., Fu Y., Hu P., Hua A., Kenton S., Lai H., Lao H.I.,
RA Lewis J., Lewis S., Lin S.-P., Loh P., Malaj E., Nguyen T., Pan H.I.,
RA Phan S., Qi S., Qian Y., Ray L., Ren Q., Shaali S., Sloan D., Song L.,
RA Wang Q., Wang Y., Wang Z., White J., Willingham D., Wu H., Yao Z.,
RA Zhan M., Zhang G., Chisoe S., Murray J., Miller N., Minx P.,
RA Fulton R., Johnson D., Bemis G., Bentley D., Bradshaw H., Bourne S.,
RA Cordes M., Du Z., Fulton L., Goela D., Graves T., Hawkins J.,
RA Hinde K., Kemp K., Latreille P., Layman D., Ozersky P., Rohlfing T.,
RA Schett P., Walker C., Wamsley A., Wohlmann P., Pepin K., Nelson J.,
RA Kort I., Bedell J.A., Hillier L., Mardis E., Waterston R., Wilson R.,
RA Emanuel B.S., Shaikh T., Kurahashi H., Salita S., Budarf M.L.,
RA Mcdermid H.E., Johnson A., Wong A.C.C., Morrow B.E., Edelmann L.,
RA Kim U.J., Shizuya H., Simon M.I., Dumanski J.P., Peyraud M., Kedra D.,
RA Seroussi E., Fransson I., Tapia I., Bruder C.E., O'Brien K.P.,
RA Wilkinson P., Bodenteich A., Hartman K., Hu X., Khan A.S., Lane L.,
RA Tiliann Y., Wright H.;
RT "The DNA sequence of human chromosome 22.";
RL Nature 402:489-495(1999).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=lymph;
RA Strausberg R.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DDAJ databases.
CC -1- FUNCTION: RESPONSIBLE FOR THE MANIFESTATION OF PERIPHERAL-TYPE
CC BENZODIAZEPINE RECOGNITION SITES AND IS MOST LIKELY TO COMPRISE
CC BINDING DOMAINS FOR BENZODIAZEPINES AND ISOQUINOLINE CARBOXAMIDES.
CC MAY PLAY A ROLE IN THE TRANSPORT OF PORPHYRINS AND HEME.
CC -1- SUBCELLULAR LOCATION: MITOCHONDRIAL; INTEGRAL MEMBRANE PROTEIN.
CC -1- SIMILARITY: SEGMENTS TMI, TM4, AND TMS SHOW SIMILARITY WITH THE

```

```

FT TRANSMEM 260 280 POTENTIAL.
FT TRANSMEM 303 323 POTENTIAL.
FT TRANSMEM 344 364 POTENTIAL.
FT TRANSMEM 377 397 POTENTIAL.
FT TRANSMEM 409 429 POTENTIAL.
FT TRANSMEM 442 462 POTENTIAL.
FT TRANSMEM 493 513 POTENTIAL.
FT TRANSMEM 528 548 POTENTIAL.
FT MOD_RES 594 594 PHOSPHORYLATION (BY SIMILARITY).
SQ SEQUENCE 674 AA; 71626 MW; 67A75AA76E42FA2 CRC64;

Query Match 2.2%; Score 11; DB 1; Length 674;
Best Local Similarity 100.0%; Pred. No. 0.042; Indels 0; Gaps 0;
Matches 11; Conservative 0; Mismatches 0;

QY 293 GAAAGGAGAA 303
DB 499 GAAAGGAGAA 509

RESULT 3
FSH_DROME STANDARD; PRT; 2038 AA.
ID FSH_DROME
AC P13709; P13710; (Rel. 13, Created)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-JAN-1990 (Rel. 13, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Female sterile homeotic protein (Fragile-chorion membrane protein).
GN FSH OR FSH.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Mandibulata; Pancrustacea; Hexapoda;
OC Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;
OC Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
OC NCBI_TaxId=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=8926730; PubMed=2567251;
RA Haynes S.R., Mozer B.A., Bhatia-Dey N., David I.B.;
RT "The Drosophila fsh locus, a maternal effect homeotic gene, encodes
RT apparent membrane proteins.";
RL Dev. Biol. 134:246-257(1989)
CC -1- FUNCTION: REQUIRED MATERIALLY FOR PROPER EXPRESSION OF OTHER
CC HOMEOTIC GENES INVOLVED IN PATTERN FORMATION, SUCH AS UBX.
CC -1- SIMILARITY: HIGH, TO HUMAN RING3 PROTEIN.
CC -1- SIMILARITY: CONTAINS 2 BROMODOMAINS.
CC -1- SIMILARITY: CONTAINS 1 ET DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M23221; AAA28540.1; -
CC EMBL; M23222; AAA28541.1; ALT TERM.
CC EMBL; M15762; AAA70424.1; -
CC EMBL; M15763; AAA70423.1; -
CC EMBL; M15764; AAA70422.1; -
CC PIR; A43742; A43742.
CC HSSP; Q92831; 1B91.
CC FlyBase; FBgn0004656; fs(1)h.
CC InterPro; IPR001487; Bromodomain.
CC Pfam; PF00439; bromodomain; 2.
CC PRINTS; PR00503; BROMODOMAIN.
CC SMART; SM00297; BROMO; 2.
CC PROSITE; PS00633; BROMODOMAIN_1; 2.
CC PROSITE; PS00144; BROMODOMAIN_2; 2.
CC Developmental Protein; Bromodomain; Transmembrane; Repeat.
KW DOMAIN 51 123 BROMODOMAIN 1.
FT DOMAIN 945 1106 ET DOMAIN.

```

```

FT TRANSMEM 330 350 POTENTIAL.
FT TRANSMEM 451 471 POTENTIAL.
FT TRANSMEM 750 770 POTENTIAL.
FT TRANSMEM 790 810 POTENTIAL.
FT TRANSMEM 816 830 POTENTIAL.
FT TRANSMEM 874 894 POTENTIAL.
FT TRANSMEM 1731 1751 POTENTIAL.
FT TRANSMEM 1939 1959 POTENTIAL.
FT VARIANT 909 909 G -> A.
FT VARIANT 1022 1022 H -> RPKY.
SQ SEQUENCE 2038 AA; 205332 MW; 849E0706D50A0098 CRC64;

Query Match 2.2%; Score 11; DB 1; Length 2038;
Best Local Similarity 100.0%; Pred. No. 0.11;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 294 AAGAGGAGAA 304
DB 343 AAGAGGAGAA 353

RESULT 4
RS21_RHIL0 STANDARD; PRT; 88 AA.
ID RS21_RHIL0
AC 098GY4;
DT 15-JUN-2002 (Rel. 41, Created)
DT 15-JUN-2002 (Rel. 41, Last sequence update)
DT 15-JUN-2002 (Rel. 41, Last annotation update)
DE 30S ribosomal protein S21.
GN RPSU OR MSR3117.
OS Ritzobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Pyllobacteriaceae; Mesorhizobium.
OC NCBI_TaxId=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAFE303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsuno M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti.";
RL DNA Res. 7:331-338(2000).
CC -1- SIMILARITY: BELONGS TO THE S21P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AP003001; BAB50082.1; -
CC InterPro; IPR001911; Ribosomal_S21.
CC Pfam; PF01165; Ribosomal_S21; 1.
CC PRINTS; PR00976; RIBOSOMAL_S21.
CC TIGRPFAM; TIGR00030; S21P; 1.
CC PROSITE; PS01181; RIBOSOMAL_S21; FALSE_NEG.
CC Ribosomal protein; Complete proteome.
KW SEQUENCE 88 AA; 10045 MW; 2C8536956632157 CRC64;

Query Match 1.8%; Score 9; DB 1; Length 88;
Best Local Similarity 100.0%; Pred. No. 0.5;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 296 AAGAGGAGAA 304
DB 76 AAGAGGAGAA 84

```

```

983 6 1.2 364 1 A3_VIGUN
984 6 1.2 364 1 FLG1_ZYMO
985 6 1.2 364 1 SYM_MERTH
986 6 1.2 365 1 MN16_HUMAN
987 6 1.2 366 1 ALF2_CAEEL
988 6 1.2 366 1 METK_PEA
989 6 1.2 367 1 ELV3_HUMAN
990 6 1.2 367 1 ELV3_MOUSE
991 6 1.2 367 1 NADA_BACHD
992 6 1.2 367 1 MECA_SALRT
993 6 1.2 367 1 MECA_SALTY
994 6 1.2 368 1 GALT_HUMAN
995 6 1.2 368 1 H181_RHIME
996 6 1.2 368 1 HIS8_AGR5
997 6 1.2 368 1 HIS8_BRUME
998 6 1.2 368 1 HIS8_BUCAT
999 6 1.2 368 1 YTL_MOUSE
1000 6 1.2 369 1 FLG1_PSEAE

```

ALIGNMENTS

```

RESULT 1
RL7_RHIME
ID 0920H8; STANDARD: PRT; 126 AA.
AC 15-JUN-2002 (rel. 41, Created)
DT 15-JUN-2002 (rel. 41, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE 50S ribosomal protein L7/L12.
DE RPL OR R01347 OR SMC01318.
OS Rhizobium meliloti (Sinorhizobium meliloti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396507; PubMed=11481430;
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreaano S., Gloux S.,
RA Godie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,
RA Pohl T., Portetelle D., Puhler A., Purnelle B., Ramsperger U.,
RA Renard C., Thebault P., Vandenbol M., Weidner S., Gallbert F.;
RT "Analysis of the chromosome sequence of the legume symbiont
RT Sinorhizobium meliloti strain 1021.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).
CC -!- FUNCTION: Seems to be the binding site for several of the factors
CC involved in protein synthesis and appears to be essential for
CC accurate translation (by similarity).
CC -!- SIMILARITY: BELONGS TO THE L12P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: AL591787; CAC45926.1; -
CC InterPro: IPR000206; Ribosomal_L12.
CC Pfam: PF00542; Ribosomal_L12; 1.
CC ProDom: PD001326; Ribosomal_L12; 1.
CC TIGRfams: TIGR00855; L12; 1.
CC Ribosomal protein; Complete proteome.
SQ SEQUENCE 126 AA; 12835 MW; 71DBF16FC487F88 CRC64;

```

```

Query Match 2.2%; Score 11; DB 1; Length 126;
Best Local Similarity 100.0%; Pred No. 0.0096;
Matches 11; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Oy 295 AAAGAGAGAAA 305
Db 41 AAAGAGAGAAA 51
RESULT 2
PTGA_CORGL
ID 045298; STANDARD: PRT; 674 AA.
AC 01-NOV-1997 (rel. 35, Created)
DT 01-NOV-1997 (rel. 35, Last sequence update)
DT 15-JUN-2002 (rel. 41, Last annotation update)
DE PTS system, glucose-specific IIBC component (EIIABC-Glc) (Glucose-
DE permease IIBC component) (Phosphotransferase enzyme II, ABC
DE component) (EC 2.7.1.69) (EII-Glc/EIIB-Glc).
GN PTSG.
OS Corynebacterium glutamicum (Brevibacterium flavum).
OG Plasmid pBSBG2.
OC Bacteria; Actinobacteria; Actinobacteria (class); Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Corynebacteriaceae;
OC Corynebacterium.
OX NCBI_TaxID=1718;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 13869;
RA Yoon K.-H.;
RT "Cloning and nucleotide sequence of enzyme II of Brevibacterium
RT lactofermentum phosphotransferase system.";
RL Submitted (NOV-1993) to the EMBL/Genbank/DBJ databases.
CC -!- FUNCTION: THIS IS A COMPONENT OF THE PHOSPHOTRANSFERASE-DEPENDENT
CC SUGAR PHOSPHOTRANSFERASE SYSTEM (PTS), A MAJOR CARBOHYDRATE ACTIVE
CC -TRANSPORT SYSTEM. THE IICD DOMAINS CONTAIN THE SUGAR BINDING SITE
CC AND THE TRANSMEMBRANE CHANNEL. THE IIA DOMAIN CONTAINS THE PRIMARY
CC PHOSPHORYLATION SITE (THE DONOR IS PHOSHO-HPR); IIB TRANSFERS ITS
CC PHOSPHORYL GROUP TO THE IIB DOMAIN WHICH FINALLY TRANSFERS IT TO
CC THE SUGAR.
CC -!- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein
CC histidine + sugar phosphate.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: CONTAINS 1 PTS EIIA DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 PTS EIIB DOMAIN.
CC -!- SIMILARITY: CONTAINS 1 PTS EIIIC DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: L18875; AAA22992.1; -
CC HSSP: P08837; IGIC.
DR InterPro: IPR001127; PTS_EIIA.
DR InterPro: IPR001996; PTS_EIIB.
DR InterPro: IPR003352; PTS_EIIC.
DR Pfam: PF00358; PTS_EIIA_1; 1.
DR Pfam: PF00367; PTS_EIIB_1.
DR Pfam: PF02378; PTS_EIIC_1.
DR ProDom: PD001476; PTS_EIIB; 1.
DR ProDom: PD002243; PTS_EIIA; 1.
DR TIGRfams: TIGR00830; PTBA; 1.
DR PROSITE: PS00371; PTS_EIIA_1; 1.
DR PROSITE: PS01035; PTS_EIIB_CYS; 1.
KW Phosphotransferase system; Sugar transport; Transferrase;
KW Phosphorylation; Transmembrane; Plasmid.
FT DOMAIN 1 43 EIIB DOMAIN.
FT DOMAIN 2 674 EIIIC DOMAIN.
FT DOMAIN 3 126 EIIA DOMAIN.
FT TRANSMEM 162 182 POTENTIAL.
FT TRANSMEM 193 213 POTENTIAL.
FT TRANSMEM 225 245 POTENTIAL.

```

837 6 1.2 315 1 COXX_BRUME
838 6 1.2 315 1 HYFC_ECOLI
839 6 1.2 315 1 SX22_HUMAN
840 6 1.2 315 1 Y060_BUCAL
841 6 1.2 317 1 FLG1_ECOLI
842 6 1.2 317 1 MTX1_HUMAN
843 6 1.2 317 1 MTX1_MOUSE
844 6 1.2 317 1 RLAO_HUMAN
845 6 1.2 317 1 RLAO_MOUSE
846 6 1.2 317 1 RLAO_RAT
847 6 1.2 317 1 SOX2_HUMAN
848 6 1.2 318 1 ADT2_YEAST
849 6 1.2 318 1 ETFA_MYCLE
850 6 1.2 318 1 ETFA_MYCTU
851 6 1.2 319 1 IFR_MEDSA
852 6 1.2 319 1 CCSA_OENHO
853 6 1.2 319 1 PSTC_MYCLE
854 6 1.2 320 1 CH36_CERCA
855 6 1.2 320 1 MP51_PHANO
856 6 1.2 320 1 SOX2_SHEEP
857 6 1.2 321 1 FCL_HUMAN
858 6 1.2 321 1 CMGB_BACSU
859 6 1.2 323 1 PE2R_RAT
860 6 1.2 323 1 TKRA_ERHME
861 6 1.2 324 1 TH14_FUSSH
862 6 1.2 325 1 A85B_MYCA
863 6 1.2 325 1 DUP3_HUMAN
864 6 1.2 325 1 ETFA_BACSU
865 6 1.2 326 1 BLAA_STRCI
866 6 1.2 326 1 TRXB_BORBU
867 6 1.2 326 1 VS09_ROTBU
868 6 1.2 327 1 ORC_SCHPO
869 6 1.2 327 1 CEBB_CHICK
870 6 1.2 328 1 MPR2_STRCO
871 6 1.2 328 1 P2Y6_HUMAN
872 6 1.2 328 1 P2Y6_RAT
873 6 1.2 328 1 PIT_BACSU
874 6 1.2 328 1 Y051_CAEEL
875 6 1.2 329 1 TECH_CHICK
876 6 1.2 329 1 A85B_MYCA
877 6 1.2 330 1 A85B_MYCIT
878 6 1.2 330 1 AR72_HUMAN
879 6 1.2 330 1 G3P1_ECOLI
880 6 1.2 330 1 G3P1_SALTY
881 6 1.2 330 1 G3PC_LEIME
882 6 1.2 330 1 ADT1_WHEAT
883 6 1.2 331 1 ADT2_WHEAT
884 6 1.2 331 1 DIV_ECOLI
885 6 1.2 331 1 MACS_HUMAN
886 6 1.2 331 1 Y181_RALSO
887 6 1.2 331 1 Y181_MOUSE
888 6 1.2 333 1 HK32_HUMAN
889 6 1.2 333 1 HME1_CHICK
890 6 1.2 333 1 ARGC_METHU
891 6 1.2 334 1 LYTE_BACSU
892 6 1.2 334 1 FABH_MYCTU
893 6 1.2 335 1 YEA6_YEAST
894 6 1.2 335 1 YG93_CLOPE
895 6 1.2 335 1 YG93_CLOPE
896 6 1.2 336 1 DLDH_ACHLA
897 6 1.2 336 1 G3P1_SCHPO
898 6 1.2 336 1 G3P1_SCHMU
899 6 1.2 336 1 G3P1_SCHMU
900 6 1.2 336 1 MRAY_CHLTR
901 6 1.2 336 1 COBU_PSEDE
902 6 1.2 337 1 G3PE_MAIZE
903 6 1.2 337 1 TH23_TRYBB
904 6 1.2 337 1 YFGA_ECOLI
905 6 1.2 337 1 YFGA_ECOLI
906 6 1.2 338 1 ETFA_MEGEL
907 6 1.2 338 1 G3P_SCHMA
908 6 1.2 338 1 YFUV_ECOLI
909 6 1.2 338 1 YFUV_ECOLI

O8Yf96 bruceia me
P77858 escherichia
O15370 homo sapien
P57168 buchnera ap
P29744 escherichia
O13505 homo sapien
P47802 mus musculu
P05588 homo sapien
P14669 mus musculu
P19945 rattus norv
P48431 homo sapien
O03241 bovine roca
P18239 saccharomyc
O33096 mycobacteri
O53275 mycobacteri
P52575 medicago sa
O9ml12 oenothera h
O50098 mycobacteri
P17110 ceratitlis c
P56164 phalaris aq
P54231 ovis aries
O13630 homo sapien
P25954 bacillus su
P21652 rattus norv
P80000 erwina her
O86344 mycobacteri
P23617 fusarium so
P21160 mycobacteri
O94586 homo sapien
P44551 bacillus su
P36651 streptomyc
P94284 borrelia bu
P30210 bovine roca
P31317 schizosacch
O05826 gallus gall
O91127 streptomyc
O15077 homo sapien
O63371 rattus norv
O34436 bacillus su
O09461 caenorhabdi
P17819 kluyveromy
P54097 gallus gall
O06947 mycobacteri
O49575 mycobacteri
O43488 homo sapien
P06877 escherichia
P24165 salmonella
O01558 leishmania
O41529 triticum ae
O41630 triticum ae
P15286 escherichia
P29966 homo sapien
O8xv50 raltosia s
O04841 mus musculu
P78367 homo sapien
O05916 gallus gall
O26934 methanobact
P54421 bacillus su
O06309 mycobacteri
O43026 schizosacch
P39953 saccharomyc
O8Yj12 clostridium
P35484 achloleplasm
P78958 schizosacch
O27652 echinococcu
O84762 chlamydia t
P28935 pseudomonas
O43247 zea mays (m
O05039 trypanosoma
P27434 escherichia
O85692 megasphaeta
P20287 schistosoma
P52137 escherichia

910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982

339 1 G3P_BRUMA
339 1 G3P_ONCVO
339 1 HEMZ_RHIME
339 1 PLSX_CLOPE
339 1 PYRD_HAEIN
339 1 PYRD_PASMU
339 1 RIBD_MYCTU
341 1 ARGC_METHA
341 1 JUND_MOUSE
341 1 JUND_RAT
342 1 HUPK_AZOVI
342 1 TRM1_THEVO
343 1 GPDA_RHILLO
343 1 HRCA_MOCGE
344 1 CD2_MOUSE
344 1 CBA6_HUMAN
344 1 COMC_METHA
344 1 MPRR_STRLI
344 1 ULB6_HCMVA
346 1 ARGQ_BACSU
346 1 PUR5_BACSU
346 1 VCS4_SYNY3
347 1 A85A_MYCAV
347 1 ADH2_YEAST
347 1 ARHC_CORGL
347 1 IDI2_LACLA
347 1 JUND_HUMAN
347 1 PUR7_HALNL
347 1 VIT3_CHICK
347 1 HOXY_AZOVI
348 1 ASG2_HAEIN
349 1 DHAS_LEPIN
349 1 DKK3_MOUSE
349 1 LEP_HAEIN
350 1 HMRO_DROME
350 1 OPSB_CONCO
351 1 HIR1_PSEAE
351 1 HRCA_MYCPN
351 1 LIP1_CANAL
351 1 TAL_NEIMA
351 1 TAL_NEIMB
351 1 VORB_METHU
352 1 EGSA_THEVO
352 1 OPSD_ZOSOP
352 1 OPSD_GOBNI
352 1 OPSD_POMTI
353 1 CHLI_CVACA
353 1 TP6A_METHU
354 1 C3X1_MOUSE
354 1 C3X1_RAT
354 1 CARX_BACST
355 1 C3X1_HUMAN
355 1 NIFS_LACDE
355 1 VCBT_ECOLI
356 1 VNS2_BTVALO
357 1 ALF_SCHPO
357 1 ALF_HAEIN
359 1 CD72_HUMAN
359 1 WNSB_MOUSE
359 1 Y199_MYCTU
359 1 METM_ACTCH
360 1 MOA2_MYCTU
360 1 VCA7_BPMU
360 1 Y572_TREPA
361 1 FDR_PICAN
361 1 Y101_NPVAC
362 1 VM25_BORHE
363 1 ADA_HUMAN

P48812 brugia mala
O01360 onchocerca
O92m50 rhizobium m
O8xyn6 clostridium
P43477 haemophilus
P57858 pasteurella
P71677 m riboflavi
O58496 methanococ
P15066 mus musculu
P52909 rattus norv
P31878 azotobacter
O97x12 thermoplas
P58142 rhizobium 1
P47447 mycoplasma
P08920 mus musculu
P40199 homo sapien
O58820 methanococ
P43161 streptomyc
P16833 human cytom
P23715 bacillus su
P12043 synochocyst
P74078 mycobacteri
O52956 saccharomyc
P00331 saccharomyc
O59279 corynebacte
O9c1f5 lactococcus
P17275 homo sapien
P17353 homo sapien
Q9hnuv halobacteri
P50600 pseudomonas
O91025 gallus gall
P40597 azotobacter
P43843 haemophilus
P41394 leptospira
O94u99 mus musculu
P44454 haemophilus
P10181 drosophila
O13227 conger cong
O9hvx0 pseudomonas
P75351 mycoplasma
O94091 candida alb
O94091 neisseria m
O94139 neisseria m
O26800 methanobact
O9h116 thermoplas
O9y422 gobbis nige
O9y7b1 thermoplas
P35403 pomatoschis
O9y9y9 zosterisess
O9tix7 cyanidum c
O27089 methanobact
O9z009 mus musculu
P35411 rattus norv
P54324 bacillus st
P49238 homo sapien
P31672 lactobacill
O05073 haemophilus
P75888 escherichia
P35065 bluetongue
P46520 schizosacch
P44429 haemophilus
P21854 homo sapien
P35366 mus musculu
P22772 mus musculu
O07723 mycobacteri
P50303 actinidia c
O53881 mycobacteri
O9t1v2 treponema p
O83582 plicha angu
P33677 autographa
P25695 orthographa
P33779 borrelia he
P00813 homo sapien


```

Db 502 AACAAAGCGTGAACTGTGTGCGACCCCTGGGTCCGGAAAGCGGTGCTCCACTGCTC 561
QY 23 LeuThrSerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyValSngluThrLys 42
Db 562 TCTGGGAAGTCAACCAAG-----AAGTCACGACGAG 591
QY 43 GlnIleGlnGlnThrArgGlnGlnLysAsnThrGluMetGluSerAspAlaThrIleAla 62
Db 592 CCCTTGGCAAAACAGCTGTCTGGCTCAGAAACTGAGAGAGAGGCAATCCCAAGCCCTC 651
QY 63 GlyAlaSerGlyLys----- 67
Db 652 GGACCCACTGCCAAGCTCGGAAGGTGTCAGCGGCCCAAGCAGAGCTCCAGTGAAGAT 711
QY 68 -----AspLysThrSerSerThrThrLysThrGluThrAlaProGlnGln 82
Db 712 TCCTCCATCTCAAGCGATGACAGATGTCGAGGTGAATTCACAGCAAAACCCAGCCGAG 771
QY 83 GlyAlaAlaAlaGlyLysGluSerSerGlu-----SerGln 94
Db 772 GCCAAGGCTTCAGACGCCCTGCAAGATCTCCAGCAAGAACAGCCCGCCCTACC 831
QY 95 LysAlaGly-----AlaAspThrGlyValSerGlyAlaAlaAlaThrThrAlaSerAsn 112
Db 832 AAGTTAGGGGAATGTGGCGCCACACCTGCTAAACACAGCGCGGCGCAGCGCTGCT 891
QY 113 ThrAlaThrLysIleAlaMetGlnThrSerIleGluGluAlaSerLysSerMetGluSer 132
Db 892 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 951
QY 133 ThrLeuGluSerLeuGlnSerLeuSerAlaAlaGlnMetLysGluValGluAla----- 150
Db 952 GACTCAGACAGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1011
QY 151 -----ValValValAlaAlaLeuSerGlyLys----- 159
Db 1012 AAAGTCCCATGTGACAGACCGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1071
QY 160 -----SerSerGlySerAlaLysLeuGluThrPro 169
Db 1072 ATCTTAGCAACCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1131
QY 170 GluLeuProLysProGlyValThrProArgSerGluValIleGluIleGluLeuAlaLeu 189
Db 1132 GAG-----AAGGACTCGGAAACCAAGCAAGTGAAGAGATTCGATGAGATGGAAG 1185
QY 190 AlaLysAlaIleGlnThrLeuGlnGlyAlaAlaThrLysSerAlaLeuSerAsnThrAlaSer 209
Db 1186 CCAGTCACTGTGAATATCTCTCAGCAAGAGACTTCTGGGAAGAGCCCTCGGCGCAGAGGT 1245
QY 210 ThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGluLysGlnAlaIleLysIle 229
Db 1246 ACCCTCAAGCCCGCCCAAGAGATCTCCAGAAAGG-----GCTCTCTCAGTCAACCCCT 1299
QY 230 AspLysGluArgGluGluGluLysGlnMetLysAlaAlaGlnIleGluLysSerLysAspLeu 249
Db 1300 GGAAGAGCAAGAGCCGTGTGTCAGCCAGCAAGGAAACCAAGAGCAAGAGCAAGTGAAG 1356
QY 250 GluGlyThrMetAspThrValAsnThrValMetIleAlaValSerValAla----- 266
Db 1357 GAGTCAGAGAGATGAGTGGGAGACACCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1416
QY 267 -----IleThrValIleSerIleValAlaAlaIlePheThrLysGlyAlaGly 282
Db 1417 AAGGTGAACCTTTGGGGAAGAGCTCCAGGTCAAGCTGTTCCACCGTCAACCCCGGGG 1476
QY 283 LeuAlaGlyLeuAlaAlaGly-----AlaAlaValAlaAlaAlaAlaAlaAlaAlaAla 297
Db 1477 TCATCGGGAAGAGGTGCCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1536
QY 297 ----- 297
Db 1537 GTCCAAATGTGAAGAAAGAGATGCTCTGAGAGCAGCAGTGCAGAGCTGGAAGTGAAC 1596

```

```

QY 298 GlyGlyAlaAlaGlyAlaAlaAlaAlaAlaThrThrValAlaAlaThrGlnIleThrValGlnAla 317
Db 1597 GGGCTGGAGGCCACAGCCCAAGGCAAGGCTCTCTGCTGCTCTCTCAGAGGTGAGGCCCT 1656
QY 318 ValValGlnAlaValLys-----GlnAlaVal 326
Db 1657 GTGGCCACCCAGCTCAAGACTCAGACGGGGCAAGGCCACTCAGGAGCAGTGAAGATCA 1716
QY 327 IleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyIle 346
Db 1717 TCTGACAGGAGAGAGAGGACACACAGCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1776
QY 347 LysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSerGlyLysIleSer 366
Db 1777 GAAAGCAGATGAAGCTTCTCTGTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1836
QY 367 LysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPheProLysLeuSerLysVal 386
Db 1837 -----ACCTGCTCCATTTGAAGCAGAGCGGTG 1866
QY 387 IleSerSerLeuThrSerLysTrpValThrValGlyValGlyValAlaValAlaAlaPro 406
Db 1867 ACCCTCTCAGCCAGC-----CTGTCATCCCA 1893
QY 407 AlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGluMetGlnGlnAsn-ValAlaGly 426
Db 1894 GCTCTGGCCAAAGGC-----ACCCAGAGCTGAGATGTGACTCTTCCAGTGAAGTCT 1944
QY 426 nPheGlnLysGluVal-----GlyLysLeuGlnAlaAla 438
Db 1945 GAGTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2004
QY 438 AspMetIleSerMet---PheThrGlnPheThrGlnGlnAla----- 451
Db 2005 CTCCAAGGGAAGAGCTCTTGGGCAAGGGGTGGCCCAAGTCAACATCAGAAAGCAGG 2064
QY 452 -----SerLysIleAlaSerLysGlnThrGlyLysSerAsnGluMetThrGluLysAl 469
Db 2065 CCTTGGGTAAAGCTATGCTCAGGAAGACTCAGAGAGCTCGAGAGAGACTCCAGACAGC 2124
QY 469 aThrLys 471
Db 2125 GAGGAAG 2131

```

RESULT 44

AZ312601

LOCUS 526 bp DNA linear GSS 29-SEP-2000

DEFINITION 1M0028104R Mouse 10kb plasmid U06C1M library Mus musculus genomic clone U06C1M0028104 R, DNA sequence.

ACCESSION AZ312601

VERSION AZ312601.1 GI:10356717

KEYWORDS GSS.

SOURCE house mouse.

ORGANISM Mus musculus.

REFERENCE 1 (bases 1 to 526)

Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C., Islam,R., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly,M., Rose,M., Rose,R., Stokes,R., Tinney,A., von Niederhausern,A. and Wright,D.,Weiss,R.

Mouse whole genome scaffolding with paired end reads from 10kb plasmid inserts

UNPUBLISHED (2000)

CONTACT: Robert B. Weiss

UNIVERSITY OF UTAH GENOME CENTER

RM. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLIC, UT 84112, USA

TEL: 801 585 5606

FAX: 801 585 7177

EMAIL: ddunn@genetics.utah.edu


```

Qy 142 AlaAlaGlnMetLysGluValAlaGluValAlaAlaAlaLeuSerGlyLysSer 161
Db 563 AAGAGCAAGCTAAGGAGGACACAGAGATGCTTCAGAGAGGACCAAGGCTGACACT 622
Qy 162 GlySerAlaLysLeuGluThrProGluLeuProLysPro----- 174
Db 623 GGGAGAGCGCAAGTAGAAGACAGCTGAGCCCGGAGCAGAGANTGCAGCTGTGCAG 682
Qy 175 -----GlyValThrProArgSerLysValIleGluIleGlyLeuAlaLeuAlaLys 191
Db 683 GCAGCGATGACGGCCGCTATGCGGACCATGTGAGGTCGACACAGCTGCAGGGCAAG 742
Qy 192 AlaIleGlnThrLeuGlyGluAlaThrLysSerAlaLeuSer----- 205
Db 743 ---ATCCGACTCTTCAGAGAACAGCTGAGAACGCTTACACCCAGCTGGCCGCTCA 799
Qy 206 -----AsnTyrAlaSerThrGlnAlaGlnAla 214
Db 800 CACGAGAGAACTCCATCTTGGAGGATGCTTGAACAGCTACACGAGCTGAGGAAAGC 859
Qy 215 AspGlnThrAsnLysLeuGly---LeuGluLysGlnAlaIleLysIleAspLysGlu--- 232
Db 860 AAGCAGAACACAGACTGCGCAAGCTCGACAGGAGCTCAGCAACAGTCAACAGAGCTA 919
Qy 233 -----ArgGluLutTyrGlnGluMetLysAlaAlaGluLys 245
Db 920 GTGGAGAGTCAAGAGCGCTCCCGGAGAGAGAGCAACAGGAGGCGCTGGAAGCCAA 979
Qy 246 SerLysAspLeuGluGlyTyrMetAspThrValAsnThrValMetIleAlaValSerVal 265
Db 980 GCAGCGACCTTGAAGAACAGAGCTCT-ACAGT----- 1011
Qy 266 AlaIleThrValIleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGly 285
Db 1012 -----GCGAGC 1017
Qy 286 Leu-----AlaAlaGlyAlaAlaValAlaAlaAlaAlaGly--- 298
Db 1018 ATCCCAACAGAGAGTAGAAGAGCTCTGCAAGAGCGCTTGGAGAGGACCCCGGACT 1077
Qy 299 -----GlyAlaAlaGly 302
Db 1078 CTGCCGGGACAAACGACCATGCCACCTCCGAGCGGATGCTGAAGGCTCAGAGACA 1137
Qy 303 AlaAlaAlaAlaThrThrValAlaThrGlnIlePheThrValGlnAlaValAlaGlnAla 322
Db 1138 GCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1197
Qy 322 LysGlnAlaValIleThrAlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaAla 342
Db 1198 CAAGTGTCAACAGCTGACTACTGCTCATGAGGAGCTCAAGAGAGGAGGAGGAGGAG 1251
Qy 342 LysSerGlyIleLysAlaPheIleLysThrLeuValLysAlaIleAlaLysAlaIleSe 362
Db 1252 GAACCTCAACGCTCAGAGAGAGAGATCCGTCATCAACAGGAGGAGGAGGAGGAG 1311
Qy 362 LysGlyLysLeuSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPhePro 382
Db 1312 CCAAGACACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1348
Qy 382 LysSerLysValIleSerSerLeuThrSerLysThrValThrValGlyValAlaVal 402
Db 1348 ----- 1348
Qy 402 ValAlaAlaProAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGluMetGln 422
Db 1349 -----CACGAGACTCGCTCAAGAGAGCTGAGCTC 1377
Qy 422 AsnValAlaGlnPheGlnLysGluValAlaGlyLysLeuGlnAlaAlaAlaAspMetIleSe 442
Db 1378 CCAGGTGCTGTCTGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1429
Qy 442 MetPheThrGlnPheThrGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyLys 462

```

```

Db 1430 -----CAACAGAGGAGGAGAA 1446
Qy 462 rAsnGluMetThrGlnLysAlaThrLys 471
Db 1447 CAATGACCTCCGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1474
RESULT 42
BC030898 3441 bp mRNA linear HTC BC030898
LOCUS Mus musculus, similar to host cell factor C1, clone IMAGE:4502186,
DEFINITION mRNA.
ACCESSION BC030898
VERSION BC030898.1 GI:21410147
KEYWORDS HTC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 3441)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (03-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NHL-MGC Project URL: http://mgc.nci.nih.gov
REMARK Contact: MGC help desk
COMMENT Email: cgaps-remail.nih.gov
Tissue Procurement: The Cepko Laboratory
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.sngc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAX Plate: 44 Row: e Column: 14
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis
This clone has the following problem: incomplete processing.
FEATURES
source
1..3441
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:4502186"
/rissue_type="Eye, retina, mouse strain C57BL/6"
/clone_id="NHL_MGC_94"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
BASE COUNT 849 a 969 c 813 g 810 t
ORIGIN
Alignment Scores:
Pred. No.: 0.183 Length: 3441
Score: 126.00 Matches: 73
Percent Similarity: 32.83% Conservative: 58
Best Local Similarity: 18.30% Mismatches: 122
Query Match: 5.32% Indels: 146
DB: 11 Gaps: 10
US-09-889-314-2 (1-496) x BC030898 (1-3441)
Qy 65 SerGlyLysAspLysThrSerSerThrThrLysThrGluThrAlaProGlnGlnGly--- 83
Db 2 ACAGGTACACCGACACACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 61
Qy 84 ---ValAlaAlaGlyLysGluSerSerGlnSerGlnLysAlaGlyAlaAspThrGlyVal 102

```

Saito, T., Okazaki, Y., Gojobori, T., Bono, H., Kasukawa, T., Saito, R., Kadota, K., Matsuda, H., Ashburner, M., Batalov, S., Casavant, T., Fleischmann, M., Gaasterland, T., Gissi, C., King, B., Kochwa, H., Kuehl, P., Lewis, S., Matsuo, Y., Nikaido, I., Pesole, G., Quackenbush, J., Schriml, L. M., Stabili, F., Suzuki, R., Tomita, M., Wagner, L., Washio, T., Sakai, K., Okido, T., Furuno, M., Aono, H., Baldarelli, R., Barsh, G., Blake, J., Boffelli, D., Bojunga, N., Carinci, P., de Bonaldo, M. F., Brownstein, M. J., Bult, C., Fletcher, C., Fujita, M., Gariboldi, M., Gustincich, S., Hill, D., Hofmann, L., Hume, D. A., Kamiya, M., Lee, N. H., Lyons, P., Mochizuki, L., Mashima, J., Mazzarelli, J., Mombarts, P., Nordone, P., Ring, B., Ringwald, M., Rodriguez, I., Sakamoto, N., Sasaki, H., Sato, K., Schonbach, C., Seta, T., Shibata, Y., Storch, K. F., Suzuki, H., Toyooka, K., Wang, K. H., Weitz, C., Whitaker, C., Wilming, L., Wynshaw-Boris, A., Yoshida, K., Hasegawa, Y., Kawaji, H., Kohsaki, S., and Hayashizaki, Y.

TITLE
Functional annotation of a full-length mouse cDNA collection

JOURNAL
Nature 409 (6821), 685-690 (2001)

MEDLINE
21085660

PUBMED
11217851

REFERENCE
5 (bases 1 to 3110)

AUTHORS
Adachi, J., Aizawa, K., Akahira, S., Akimura, T., Aono, H., Arai, A., Arakawa, T., Baldarelli, R., Bono, H., Brownstein, M., Bult, C., Carinci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hasegawa, T., Hara, A., Hayatsu, N., Hill, D., Hiramoto, K., Hirooka, T., Hori, F., Hume, D., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Quackenbush, J., Saito, H., Saito, R., Sakai, C., Sakai, C., Sano, H., Sasaki, D., Schriml, L., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sobabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, T., Tanaka, T., Tejima, Y., Toya, T., Yamamura, T., Yamana, I., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.

TITLE
Direct Submission

JOURNAL
Submitted (18-AUG-2000) Yoshinide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-resgsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5'] GAGAGAGAGAGATCCAGATCCAGATCTTTTTTTTTTTTNN 3'], cDNA was prepared by using trehalose thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5'] GAGAGAGATCTCGAGTAAATTAATATATCCCCCCCC 3'. cDNA was cleaved with BamHI and XhoI. Vector: a modified phlucscript KS(+) after bulk excision from lambda FLC I. Cloning sites, 5' end: SalI, 3' end: BamHI. Host: DH10B.

FEATURES
source
1. 3110
Location/Qualifiers
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="PANTOM_DB:5730465C04"
/db_xref="MGD:MGI:1912849"
/db_xref="taxon:10090"
/clone="5730465C04"
/clone.lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="8 days embryo"
1..2614
/note="data source:SPRR, source key:075300, evidence:ISS homolog to ES/130 putative"

CDS

```

/codon_start=2
/protein_id="BAB31939.1"
/db_xref="GI:12860388"
/db_xref="MGD:MGI:1926071"
/translation="ASNONRTDVAANGTQTOEGVSNOYKSESSPNQKAEAPNO
GKKKDSPOAKRYDAANQKSEMAAOQKSMVSOAPQADAPAKKSSRRK
GEPGPPDDGLFLPYKTLVSTVSMSFSEBAQRLIEISEKGVIODVMKKTQK
DPAVILKROLEKEKELATEQDAVAASKRLINKEMASERKAAGAAGAAKQOLA
REOETAAARMQASRYRDHVEVOLOQLEITLOEONGENTOLARJOENSLIRDA
LNQATSOYESKONMETLAKUROLBSKVNEIYKSPASQEPQORALFAKATFEQV
LOLQASHSESEDAKRLIEYRELICRAQTSHANIRADAEKAOQOQVATLHSKLOS
SEVEKSKCEBLSLHGLKBARENSOQLTERINSIEALLAQAQDQASHAAENQO
QTRKLEESQVSCLEKETSEIKEMAEQOKNNDRKNEMKAMALALAEVAKCQTLR
SLTOAKSESEKOLHIAEQOTKEITALLPGISISAHONYAEMLOEFKEGSELLKP
TLPEMDIVLKLREAEQNTLOAEQDQRTILLETBGMLOKSVSEEEVWKAVY
GAEBELHKSRYTVYKHLEDIYKELKELESDDQVREHSHLEAELEKMAAASAEON
YAKEVAGLRLOLESQSDAEKQOSBELALVROOLSDMSHVEDGVYASPAY
PPAEQDPKRLQLETEATLEAQETROKRLAEFEERQTRACRLOEELKRAAGPL
ESSGKEITQLEKLEKRLKRLSDIGRAAIRLOELTKTQDOLTKENDTVKLEQLEQ
KAEDSSSEKSTSV"
3087..3092
polyA_signal
/note="putative"
3110
polyA_site
/note="putative"
BASE COUNT 930 a 771 c 936 g 473 t
ORIGIN
Alignment Scores:
Pred. No.: 0.158 Length: 3110
Score: 126.00 Matches: 98
Percent Similarity: 32.98% Conservative: 90
Best Local Similarity: 17.19% Mismatches: 176
Query Match: 5.32% Gaps: 207
DB: 11 Indels: 17

US-09-889-314-2 (1-496) x AK019964 (1-3110)
QY 6 ILeSerSerSerSerGlyProAsnGlnLysAsnIleMetSerGlnValLeuThrSer 25
Db 74 GTCAAAAGTCAAGAGGTCCTCATATCA-----GGTAA 109
QY 26 ThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnThrLysGlnIleGln 45
Db 110 AAGGCAGAGGAGCCCAACCAAGCAAGCAAGAAAGATGATCCCGCCAGCCAGCCAA 169
QY 46 Gln-----ThrrgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIle 61
Db 170 AAGGTGATGCACGTCCCAATCAGGTTAAAGTCAAGATG----- 211
QY 62 AlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGln 81
Db 212 -----GCTCTCCCAAGGTCATAAAGCAAGCATGTGTCAGAGCAAGGACCAAG 265
QY 82 GlnGlyValAlaAlaGlyLysGlnSerSerGlnLysAla----- 96
Db 266 CAAGAGCAGCACTGCATAAAGAAAGATCTGTTCCAGCAAAAAGAGAGCCCTGGCCCCA 325
QY 96 ----- 96
Db 326 GATTGCGAGGGCCCTCTCTCTGCTCCCTACAGACACTGCTCTACGCTTGGAACATG 385
QY 97 -----GlyAlaAspThrGlnVal 102
Db 386 GTCTTCAGTGAAGGCGAGCGCCAGCGGCTTATTGACATTTCTCTGAGAAGACTGAGATC 445
QY 103 SerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAla---MetGlnThr 121
Db 446 ATTCAGACACCTCGCATGATGACCTCAGAAAGGTCGACCTGTACCATTCATTCGAACCC 505
QY 122 SerTleGlnGlnLalaSerLysSerMetGlnSerThrLengLysSerLeuGlnSerLeuSer 141
Db 506 CAACTCAAGAGAAAGAAAGCTGTGCGCCAGACAGACAGAGAT--GCGGCGGTGCGCC 562

```


Library="this sequence is part of a project of ESR assemblies resulting from the application of public configs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the ongoing addressing of BACs in conjunction with the Maize Mapping Project"

```

      /db_xref="MGD:MG1:1903394"
      /db_xref="Eaxon:10090"
      /clone="4921511116"
      /sex="male"
      /tissue_type="testis"
      /clone_lib="RIKEN full-length enriched mouse cDNA library"
      /dev_stage="adult"
      /note="data source:SPTR, source key:Q9NTH0, evidence:ISS
      homolog to PUTATIVE SERINE-RICH PROTEIN (FRAGMENT)"
      /db_xref="MGD:MG1:1918175"
BASE COUNT   475 a   346 c   391 g   378 t
ORIGIN

Alignment Scores:
Pred. No.:      0.0546      Length:      1590
Score:          126.50      Matches:      93
Percent Similarity: 39.78%      Conservative: 90
Best Local Similarity: 20.22%      Mismatches: 195
Query Match:      5.34%      Indels:      84
DB:               11      Gaps:      17

US-09-889-314-2 (1-496) x AK014874 (1-1590)
QY   23 LeuThrSerThrProGlnGlyValProGlnGlnAspLys-----LeuSerGlyAsn 39
      ::::::::::: ::::::::::: :::::::::::
DB   74 ATATCAATGATGACATGAGGGGTAGAGAGCCATCCAAAAGAGAGCCGCTAGAGAGAAA 133
      ::::::::::: ::::::::::: :::::::::::
QY   40 GluThrLysGlnIleGlnInThrArgGlnGlyLysAsnThrGlnMetGluSerAspAla 59
      ::::::::::: ::::::::::: :::::::::::
DB   134 AACAAATCTTCACGTGAGGAGATCATGGGAAAAATCTGCCAAAAGCATGCTTCAGCA 193
      ::::::::::: ::::::::::: :::::::::::
QY   60 ThrIleAlaGlyAlaSerGlyLysAspLysThrSerSerThrLysThrGlnThrAla 79
      ::::::::::: ::::::::::: :::::::::::
DB   194 GCTCAGCAAGAAAACAGATGCGCGCTCAGGGGGTCTCTCC--ACCAAAATCCGAGAGTGT 250
      ::::::::::: ::::::::::: :::::::::::
QY   80 ProGlnGlnGlyAlaAlaGlyLysGluSerSerGluSer---GlnLysAlaGlyAla 98
      ::::::::::: ::::::::::: :::::::::::
DB   251 -----GGGACCTGAGTCTCGAGCAGCTCCAGTGGCTCCCATCAGCATGATGAGCCACA 301
      ::::::::::: ::::::::::: :::::::::::
QY   99 AspThrGlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAla 118
      ::::::::::: ::::::::::: :::::::::::
DB   302 AGTGAAGGAGACAGATCTGTGTGACAGCAGAGCAGCAATTCCTCTCAGTAAACCGCA 361
      ::::::::::: ::::::::::: :::::::::::
QY   119 -----MetGlnThrSerIleGlnGluAla 126
      ::::::::::: ::::::::::: :::::::::::
DB   362 GGGCTCGAAGAGCTTGGAGTCAAGATCCACATAAACGTGCTCAGGCTCATTTGTT 421
      ::::::::::: ::::::::::: :::::::::::
QY   127 SerLysSerMetGluSerThrLysGluSerLysGlnSerLysSerAlaIleGlnMetLys 146
      ::::::::::: ::::::::::: :::::::::::
DB   422 TCTTCCCTGTTGAAGTCCATATGATATGATACCCAGTCCACTGATTAACAGACAA 481
      ::::::::::: ::::::::::: :::::::::::
QY   147 GluValGluAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlySerAlaLysLeu 166
      ::::::::::: ::::::::::: :::::::::::
DB   482 CAGAGGGGGTCACTAAAGAGGGGCTCTGCAAGGT---TCTCAGGCTCAGCCTCTCAG 538
      ::::::::::: ::::::::::: :::::::::::
QY   167 GluThrProGluLeuProLysProGlyValThrProArgSerGluValIleGlnIleGly 186
      ::::::::::: ::::::::::: :::::::::::
DB   539 AGCAGCTCAAGATTTGAGTGGCTTGTGTGCTCT---CTGGAAGTGCA-GAAGTAGAG 594
      ::::::::::: ::::::::::: :::::::::::
QY   187 LeuAlaLeuAlaLysAlaIleGlnThrLysGlnGluAlaThrLysSerAlaLeuSerAsn 206
      ::::::::::: ::::::::::: :::::::::::
DB   595 TTGGCATTTACTG---TCTTGAAGTCAAGTTCAAGACAGACAGTTCAGTGGTTAACTCT 651
      ::::::::::: ::::::::::: :::::::::::
QY   207 TyrAlaSerThrGlnAlaGlnAlaAlaAspGlnThrAsnLysLeuGlnGluLysGlnAla 226
      ::::::::::: ::::::::::: :::::::::::
DB   652 AAATCTAGTTCAGAGCAAAACATTTCAATCATCA----- 684
      ::::::::::: ::::::::::: :::::::::::
QY   227 IleLysIleAspLysGluValGlnGluThrGlnGlnMetLysAlaIleGlnGlnLysSer 246
      ::::::::::: ::::::::::: :::::::::::
DB   685 -----GTTTCAAAAACAGTTTCATGCTCAGGAGCTCTTACTAACCCCCCAGAGAGT 738
      ::::::::::: ::::::::::: :::::::::::
QY   247 LysAspLeuGlnGlyThrMetAspThrValAsnThrValMetIleAlaValSerValAla 266

```


Oy 100 ThrcgYvalSerGlyAlaAlaAlaAlaThrThrLa---SeranThrAlaThrLySIIAla 118
 Db 248 --GGGACCCCTGGGAAAGGGGCTACCCACACCCCTGGGAAAGCAGGGCGTGTACC 304
 Oy 119 MetGlnThrSerIleGluGluAlaSerLySerMetGlnSerThrLeuGln 138
 Db 305 TCCGAGACCAAGCCAGGAAACCCAGAGAGACATCTCAGACAGACAGACAGAGAG---TCA 361
 Oy 139 SerLeuSerAlaAlaGlnMetLySgluValAlaValAlaValAlaAlaLeuSerGly 158
 Db 362 TCTGCACCTGAGGAGAAACCCAGCTCCCAAGGCCCTTCAGCGCAAGCCCTCAGA 421
 Oy 159 LySerSer-----GlySerAlaLySleuGluThrProGluLeuProLySProGlyVal 176
 Db 422 AAAACCTCTCAGAGTCGGAGCTGCTCCAGCCCTGCCAAGAGACTCCCCAGAAAGAGACT 481
 Oy 177 ThrProCysSerGluValIleGluIleGlyLeuAlaLeuAlaAlaIleGlnThrLeu 196
 Db 482 GCCCA---GGCCCCCTGGGAAGACAGAGGCGCTGCAGTGGCCAAAGCCCAAGGGGAG 538
 Oy 197 GlyAlaAlaThrLySerSerAlaLeuSerAnThrAlaSerThrGln-----Ala 212
 Db 539 CGGAGGAGAGACTCCAGACAGCAGCAGAGATCGCAGTGAAGAGAGAGAGCGCTGCT 598
 Oy 213 GlnAlaSerGlnThrAsnLySleuGly----- 221
 Db 599 CAGGGAGAACCTTCAGGAAAGCCCCCAGCTCAGAGCCGCTCGGCCCTGCCAAGAG 658
 Oy 222 -----LeuGlnLySglnAlaIleLySIIeAspLyS 231
 Db 659 TCCCCCAGAAAGGGCTGCCCCAGCAACCTCTAGAGAAAACAGGGCTCGAGCGGCCAG 718
 Oy 232 GlnArg-----GluGluThrGlnGluMetLySIIaAlaGluGlnLySerLySasp 248
 Db 719 TCCAGGTGGGGAGACAGAGAGAGACTCAGAAAGCAGCAGCAGAGAGTCAACAGTAC 778
 Oy 249 LeuGluGlyThr 252
 Db 779 AGAAGAGGCACT 790

RESULT 30
 FR0006944/c
 LOCUS F.rubripes GSS sequence, clone 133B1aC7, genomic survey sequence.
 DEFINITION
 ACCESSION Z90754
 VERSION Z90754.1 GI:1867968
 KEYWORDS GSS; genome survey sequence.
 SOURCE Takifugu rubripes.
 ORGANISM Takifugu rubripes.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorphia; Acanthopterygii; Percomorphia; Tetraodontiformes;
 Tetraodontidae; Takifugu.
 1 (bases 1 to 619)
 Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umanian,Y.,
 Williams,G. and Brenner,S.
 Direct Submission
 Submitted (18-FEB-1997) MRC Human Genome Mapping Project Resource
 Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hmp.mrc.ac.uk
 Vector: pBluescript II KS
 V-type: phagemid
 PRIMER: KS
 DESCR: One pass dye-terminator sequencing of cosmid cloned genomic
 sequence.

Alignment Scores:	
pred. No.:	0.0102
Score:	128.00
Percent Similarity:	49.22%
Best Local Similarity:	30.47%
Query Match:	5.40%
DB:	17
Length:	619
Matches:	39
Conservative:	24
Mismatches:	61
Indels:	4
Gaps:	3

US-09-889-314-2 (1-496) x FR0006944 (1-619)

Qy	266	AlaIleThrValIIleSerTlleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGly	285
		:::::: :::::	
Db	373	TCAATAGACAGTAGTAGTCAGCACGACACACACAGTACANTAGTACGACGACAGTAGCA---	317
Qy	286	LeuAlaIleAglyValAlaIleValGlyAlaIleAlaIleAglyGlyAlaAlaGlyValAlaAla	305
		::::::::::	
Db	316	GTAAGACAGCATAGTACGTAGTACCTAGTACGACGACA-----GCAGCAGCAGCAGCAGCA	263
Qy	306	AlaThrTrhValAlaIleThcGlnIleThrcValGlnAlaIleValAlaGlnAlaValLysGlnAla	325
		::::	
Db	262	GTAAGTAGTAGTAGTAGTAGCAGCAGCATAGTACAGTAGTACGACAGCAGCATAGTACGACAGA	203
Qy	326	ValIleThrAlaValArgGlnAlaIleThrAlaIleAlaIleValAlaValLysSerGly	345
		: ::: :: :::	
Db	202	GTAAGTAGTAGCAGTACGACGACGACATAGCAGCAGCAGCATAGTACGAGTACGACACAGCA	143
Qy	346	IleLysAlaPheIleLysThrIleuValLysAlaIleAlaIleAlaIleSerLysGlyIle	365
		:::: :: : ::: :	
Db	142	GCAGCAGCGACGACAGTACTAGTACGTACGACGATAGCA--GTAGTACACACAGTACGA	86
Qy	366	SerLysValIPheAlaLysGlyThrGlnMetIleAlaLysAsnPhcProLysLeuSerLys	385
		::: ::::: : ::: :	
Db	85	GCAGTAGTGACGACGATAGCAGCAGTACGACATTAGTACGAGNMGCTAGTACGCTAGCAGCA	26
Qy	386	ValIleSerSerLeuThrSerLys	393
		:::::	
Db	25	GTNNTAGTAGCAGTCCCTTNGAAG	2
RESULT 31			
BM58321/C			
LOCUS	BM58321	711 bp	mRNA linear EST 25-FEB-2002
DEFINITION	17000687322518 A.Gam.ad.cDNA.blood1 Anopheles gambiae cDNA clone		
ACCESSION	19600449697628 5', mRNA sequence.		
VERSION	BM58321		
KEYWORDS	BM58321.1 GI:18884182		
SOURCE	EST.		
ORGANISM	African malaria mosquito.		
	Anopheles gambiae		
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
	Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;		
	Anopheles.		
REFERENCE	1 (bases 1 to 711)		
AUTHORS	Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab		
TITLE	R.,Collins,F.H., Venter,U.C. and Hoffman,S.L.		
JOURNAL	Celera Anopheles gambiae EST project		
COMMENT	Unpublished (2002)		
	Contact: Holt R.A.		
	Celera Genomics		
	45 W. Gude Dr., Rockville, MD 20850, USA		
	Tel: 2404533151		
	Fax: 2404534580		
	Email: HoltRA@celera.com		
	Plate: NU01004AAR row: F column: 14		
	Seq primer: M13 Reverse		

FEATURES	Location/Qualifiers
SOURCE	1. 619 /organism="Takifugu rubripes" /db_xref="taxon:31033" /clone="133B16aC7" /clone_lib="cosmid 133B16"
BASE COUNT	75 a 150 c 116 g 175 t 103 others
ORIGIN	

FEATURES	SOURCE	LOCATION/Qualifiers
1.	711	
/organism="Anopheles gambiae"		
/strain="RSP-St (Reduced susc.		to Permethrin - std.
chromosome)"		
/db_xref="taxon:7165"		
/clone="19600449697628"		
/clone_lib="A.Gam.ad.cDNA.blood1"		

US-09-889-314-2 (1-496) x BG696282 (1-751)	Alignment Scores:	Pred. No.:	Score:	Percent Similarity:	Best local Similarity:	Query Match:	DB:
119 MetGlnThrSerIleGluGlnAlaSerLysSerMetGluSerThrIleuGlnSerIleuGln	0.0119	Length:	751				
19 GTCAGGCCCGCATGACGAGATATATGCCAGAGAGCAAAAGCTGAGCGCCGCTCTGTAC	128.50	Matches:	58	42.13%	109	5.42%	12
139 SerLeuSerAlaAlaGlnMetLysGluGlnAlaValAlaValAlaAlaLeuSerGly	128.50	Conservative:	49	42.13%	109	5.42%	12
79 -----CACAGCAAGTATGAAAGCTGCAGATCATCTGCGACACATGGG	128.50	Mismatches:	38	42.13%	109	5.42%	12
159 LysSerSerGlySerAlaLysLysGlnThrProGluLeuProLysProGlyValThrPro	128.50	Indels:	38	42.13%	109	5.42%	12
124 GATACCTGAGCAAAATTCAGATATGAAATTTCTGAGCTGAATGTCATCAGACAGACTT	128.50	Gaps:	6	42.13%	109	5.42%	12
179 ArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGlu	128.50						
184 AGATCTGAAATCGCAATGTCAGAACAGACATCTCCAACTTGCGACAGTCCATCAGATGAT	128.50						
199 AlaThrLysSerAlaLeuSerAsnThrLysSerThrGlnAlaGlnAlaAspGlnThrAsn	128.50						
244 GCAGAGCAG-----CGTGGCAGATATGCCCTCAAGGATGCCAAGAAC	128.50						
219 LysLeuGlyLeuGlnLysGlnAlaIleLysLysIleAspLysGlu	128.50						
286 AGCTGATGATGACCTGGAGATGGCCCTGACGACGACCAAGGAACCTGGCCGCTGCTG	128.50						
234 GlnGluThrGlnGlnMetLysAlaAlaGlu-----	128.50						
346 CGTGACTACACAGGACCTGATGACACACACAGCTGGCCCTGATCTGAGATGGCACCTAC	128.50						
244 -----GlnLysSerLysAspLeuGlnLysThrMetAspThrValAsnThrValMet	128.50						
406 AGGACCCCTCTGGAGGAGGAAGAATTCAGATGTCGTGAGACATATGCCCCGAAACGTAGT	128.50						
261 IleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCysGly	128.50						
466 GGTCTGTGAGCAACAGCACACACCATTCAGTGGAGGTGGCACCCAGAGGAGGGCGG	128.50						
281 AlaGlyLeu-----AlaGlyLeuAlaAlaGlyAlaAlaValAlaGlyAlaAlaGly	128.50						
526 GGTGTCTACGCTCTGGAGAGAGCAGCTATGGCTCCGAGGTGTACTATGTGTGGGA	128.50						
299 Gly-AlaAlaGlyAlaAlaAlaAla-----ThrThrValAlaThrGln	128.50						
586 GGTGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG	128.50						
312 nleThrValGlnAlaValAlaGlnAlaValAlaLysGlnAlaValIleThrAlaValArgI	128.50						
646 GGTGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG	128.50						
332 nAlaIleThrAlaAlaIleLysAlaAlaValLysSerIle	128.50						
706 GGTCTGTGAGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGGCGGG	128.50						

[illegible]


```

Db 137 CAGCCACCTGCCAGTGCAGAGAGACAGACAGATCGGAGCAG-----GCTGATGATCT 190
Qy 89 GUserSerGluSerGlnLysAlaGlyAlaAspThr-----GlyVal 102
Db 191 ATCCAAACACAGCAGATAGTACCGGCTCTGCTCCCTCCCTTACGATAGTCTGTGA 250
Qy 103 SerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSer 122
Db 251 GCTGCCGCCGATCCACTCTCCACACACACACTGGGATGCCACTGCTGCCGACAGAGA 310
Qy 123 IleGlnGluAlaSerLysSerMetGluSerThrLeuGlnSerLeuSerLeuSerAla 142
Db 311 GTTAACTGTAAACACACTCCACACACACTCCGACCGCTGCTGCCAAGCTTCTCTCT 370
Qy 143 AlaGlnMetLysGluValAlaValAlaValAlaValAlaValAlaValSerGlyLysSerSerGly 162
Db 371 -----ATGACCGAACCCTCTCTCAAGTTGCTGCA-----CAGGGA 406
Qy 163 SerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGluVal 182
Db 407 TCT-----GAGCGCTTCTTCTGCCAGAGCAGTCAGATGCTCCGTTTCCACAGAT 457
Qy 183 IleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyGluAlaThrLysSer 202
Db 458 ATT-----CTTCAACCCCGGAGAGACATGACAGACAGACAGACAGACACTTCCAGAG 508
Qy 203 AlaLeuSerAsn-----TyrAlaSerThrGlnAla 212
Db 509 CCGATGATAGTGCAGTGCAGACACACATGATGATGATGATGATGATGATGATGATGAT 568
Qy 213 GlnAlaAspGlnThrAsnLysLeuGlnLysGlnAlaIleLysIleAspLysGln 232
Db 569 GAGTCACTCAAGATGCTTCACAGCAGCTGCTGAACAACATGCGAGTA-----GACAGC 619
Qy 233 ArgGlnGluTyrGlnGlnMetLysAlaIleGlnLysSerLysAspLeuGlnLysThr 252
Db 620 TTATCTCTCTCTCAAGCTTATGCGAAGAGCAG-----ACT 658
Qy 253 MetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIle 272
Db 659 GCAACACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 706
Qy 273 ValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaIleAlaValAla 292
Db 706 ----- 706
Qy 293 GlyAlaAlaAlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 312
Db 707 -----GCCGACGACAGACAGCAGCTCAGCAGCAGCAGCAGCAGCAGCAGCAGCA 757
Qy 313 IleThrValGlnAlaValAlaValAlaValLysGlnAlaValIleThrAla 329
Db 758 TTGGCTATCCCGCTGCTTACTACAGCAGCTCAGCAAGCAGTATGCGCCGCA 808
RESULT 26
LOCUS BC007388 1745 bp mRNA Linear HTC 12-JUL-2001
DEFINITION Homo sapiens, clone IMAGE:3830953, mRNA.
ACCESSION BC007388
VERSION BC007388.1 GI:14712402
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1745)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov

```

```

COMMENT
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: ATCC/DCTP/DRP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbiology.org
contact: amadansystemsbiology.org
Anup Madan, Rachel Dickhoff, Jessica Fahey, Stephanie Ford, Julia
Greene, Mark Ketterman and Anuradha Madan

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IFAL Plate: 24 Row: d Column: 4
This clone has the following problem: frame shifted.

FEATURES
Source
Location/Qualifiers
1..1745
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3830953"
/tissue_type="Skin, melanotic melanoma."
/clone_lib="NIH_MGC_20"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

BASE COUNT 500 a 534 c 370 g 341 t
ORIGIN

Alignment Scores:
Pred. No.: 0.035 Length: 1745
Score: 129.00 Matches: 93
Percent Similarity: 37.47% Conservative: 70
Best Local Similarity: 21.38% Mismatches: 190
Query Match: 5.45% Indels: 82
DB: 11 Gaps: 18

US-09-889-314-2 (1-496) x BC007388 (1-1745)
Qy 22 ValLeuThrSerThrProGlnGlnAlaProGlnAlaAspLysLeuSerGly---AsnGln 40
Db 352 ATCATCTCAGACAGCCCGCAGGCG-----CAGCAGGGCTCTCTCCAGCGCAAAATCTT 405
Qy 41 ThrLysGlnIleGlnLysThrArgGlnGlnLysAsnThrGlnMetGluSerAspAlaThr 60
Db 406 CTAAAGCACTTCTCTCAAGCAGCAAGCAGCAAGCAGCAAGCAGCAAGCAGCAAGCAGCA 465
Qy 61 IleAlaGly---AlaSerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAla 79
Db 466 CTCACCTCCACAGCAGCAAGCAGCAAGCAGCAAGCAGCAAGCAGCAAGCAGCAAGCAGCTT 525
Qy 80 ProGlnGlnGlyValAlaAlaGlyLysGluSerSerGlnLysAlaGlyAlaAsp 99
Db 526 CCACAGAGCAGATCAACACCAAGCGAATGATGATCTCCAGCTTGAGAGCGCCGATGAC 585
Qy 526 CCACAGAGCAGATCAACACCAAGCGAATGATGATCTCCAGCTTGAGAGCGCCGATGAC 585
Qy 100 ThrGlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMet 119
Db 586 CTYAGAGAGCTTGAGCAGCTTCCCAAGACCTTCAAAACAAGCAGCAATGCAACTGATTC 645
Qy 120 ---GlnThrSerIle-----GlnGluAlaSerLysSer 129
Db 646 ACTCAGAGGTGATGGGCTCGCTATGGGGAATATATGAAATGATGATGATGATGATGATGAT 705
Qy 130 MetGluSerThrLeuGlnSerLeuGlnSerLeuSerAlaAlaGlnMetLysGluValGlu 149
Db 706 ACCATCTCTCGATTGAAGCCTTG---AACCTCAGCTTTAAGACATGTGCAAGTTGAAG 762
Qy 150 AlaValAlaValAlaAlaLeuSerGlyLysSer-----SerGlySerAlaLysLeuGlu 167
Db 763 CCACCTTTAGAGAACTGCTAATATGATCAGAGAACTCTCATCTGATTCGCTCCCTCC 822
Qy 168 ThrProGlu---LeuProLysProGlyValThrProArgSerGluValIleGluIleGly 186
Db 823 AGCCCAAGTGCCCTGAATTCAGAGAAATGAG-----GAC 858

```

BASE COUNT 318 a 313 c 257 g 204 t 9 others
 ALIGNMENT SCORES:
 Pred. No.: 0.0129 Length: 1101
 Score: 130.50 Matches: 74
 Percent Similarity: 37.78% Conservative: 45
 Best Local Similarity: 23.49% Mismatches: 159
 Query Match: 5.51% Indels: 37
 DB: Gaps: 8

US-09-889-314-2 (1-496) x CNS05HLZ (1-1101)

QY 60 ThrTleAlaIleGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGluThrAla 79
 Db 62 ACCATTTCAGAGAGCGATTCCAGAGCGAGGCTCCAGCCAAAGCTGCCGGAAAGTGTCT 121
 QY 80 ProGInGInGlyValAlaIleGlyLys----- 88
 Db 122 GCAGCCAAACCTGCAGCTCTCTCCAAAGCTCCCGCAAAGAACAGAGAGCTGTGAG 181
 QY 89 -----GluSerSerGluSerGlnLysAlaGlyAlaAspThrGlyValSerGlyAla 105
 Db 182 TCAGAGAGACTCCTCCTCCTGAG 241
 QY 106 AlaAlaIleThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGluGlu 125
 Db 242 GCCCTGCTCTCAAAATCTCTCGAGCAAGGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 301
 QY 126 AlaSerLysSerMetGluSerThrLeuGluSerLeuGlnSerLeuSerAlaIleGlnMet 145
 Db 302 AGCAGCTCAGATCCAGATCCTCCGAGATGAGAAACCGGCTTAAGAGAGAGAGAGAGAGAG 361
 QY 146 LysGluValAlaValAlaValAlaValAlaValAlaValAlaValAlaValAlaValAla 165
 Db 362 GTAAAAAACCTGCAGCAAG 418
 QY 166 LeuGluThrProGlnLeuProLysProGlyValThrProArgSerGluValIleGluIle 185
 Db 419 TCTAGTATGATCCGCGCAAGAAAGAACCC---GTCTCTCTGCGCAAGAGAGAGAGAGAG 466
 QY 186 GlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlyLys---AlaThrLysSerAlaLeu 204
 Db 467 -----GCTGCGAGCTAAAGCTGCTCTGCTGCGAGCAAGAGAGAGAGAGAGAGAGAG 520
 QY 205 SerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGluLys 224
 Db 521 GATTGAGACTCCTCTCAGAGATGAAGAACCTAAAGCTAAAGTAAACAGCTAGTAAAG 580
 QY 225 GlnAlaIleLysIleAspLysGluArgGluGluThrGlnGluMetLysAlaIle----- 242
 Db 581 CCTGCTACTTAAGGCTGCGAGCTAAAGCTGCTCTGCTGCGAGCAAGAGAGAGAGAGAGAG 640
 QY 243 -----GluGlnLysSerLysAspLeuGluGluThrMetAspThrVal 256
 Db 641 GCAGAGAGCTGATTCAGAGCTCTCTCTGAGAGATGAAGAACCTAAAGTAAAGTGA 700
 QY 257 AsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleValAlaIle 276
 Db 701 ACCAATTAACACAGCTACTCTGCGAGTAAAGCTACTCTGCGCTAAGGCTGCGCTG 760
 QY 277 PhetThrCysGlyAlaGlyLeuAlaGlyLeuAlaGlyAlaValAlaValAlaValAla 296
 Db 761 -----GCTAAAGCTGGTACTCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 805
 QY 297 AlaGlyValAlaAlaGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 314
 Db 806 AAAGCTGCTGCTGAG 865
 QY 315 ValGlnAlaValAlaValAlaValLysGlnAlaValIleThrAlaValAlaArgAlaIle 334
 Db 866 CCCAAAGCTTAAGCAACAGCGGCTAAAGCTGCTAAGGCTGCTGAGAGCTGCTAAGAGCTGCT 925

QY 335 ThrAlaAlaIleLysAlaAlaValLysSerGlyLysIleLysAlaPhe 349
 Db 926 CCTCTGCGAGAGAGCTGCTGCTGCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 970

RESULT 25
 B0735450 887 bp mRNA linear EST 16-JUL-2002
 LOCUS AGENCOURT 8148898 NICH04 XGC Emb4 Xenopus laevis cDNA clone
 DEFINITION IMAGE:5570055.5', mRNA sequence.
 ACCESSION B0735450.1 GI:21874347
 VERSION B0735450.1
 KEYWORDS EST.
 SOURCE African clawed frog.
 ORGANISM Xenopus laevis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 Xenopodinae; Xenopus.
 1 (bases 1 to 887)
 REFERENCE NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.
 AUTHORS National Cancer Institute, Cancer Genome Anatomy Project (CGAP).
 TITLE Tumor Gene Index
 JOURNAL Unpublished (1997)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Dr. Igor David
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
www.bio.llnl.gov/biopr/image/image.html
 Plate: LLAM12311 row: k column: 16
 High quality sequence stop: 666.
 Location/Qualifiers
 source
 1..887
 /organism="Xenopus laevis"
 /db_xref="taxon:8355"
 /clone_image="5570055"
 /clone_lib="NICH04 XGC Emb4"
 /dev_stage="embryo, stage 31-32"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: whole embryo; Vector: pCMV-SPORT6; Site: 1
 NotI; Site: 2: SalI; Cloned unidirectionally. Primer: Oligo
 dT. Average insert size 2.1 kb. Constructed by Life
 Technologies. Note: This is a Xenopus Gene Collection (XGC
) library."
 BASE COUNT 238 a 273 c 212 g 164 t
 ORIGIN

ALIGNMENT SCORES:
 Pred. No.: 0.0107 Length: 887
 Score: 130.00 Matches: 76
 Percent Similarity: 37.69% Conservative: 51
 Best Local Similarity: 22.55% Mismatches: 126
 Query Match: 5.49% Indels: 84
 DB: Gaps: 12

US-09-889-314-2 (1-496) x B0735450 (1-887)

QY 9 SerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeuThrSerThrProGln 28
 Db 2 ACATCTGCGATAGCAACAG-----CCTACT 28
 QY 29 GlyAlaProGlnGlnAspLysLeuSerGlyAsnGluThrLysGlnIleGlnInThrArg 48
 Db 29 GCTACTCCAACTCAGCAAGTCTCTCCACACCTCTTGT-----GAACTCAC 76
 QY 49 GInGlyLysAsnThrGluMetGluSerAspAlaThrIleAlaGlyAlaSerGlyLysAsp 68
 Db 77 GAGAGTGGCAACCCATAGCCAGCCAGCAAGATTAAGAGAGAGAGAGAGAGAGAGAGAGAG 136
 QY 69 LysThrSerThrThrLysThrGluThrAlaProGlnGlnGlyValAlaAlaGlyLys 88

TITLE Genomic exploration of the hemiascomycetous yeasts: 11.
 Kluveromyces lactis
 JOURNAL FEMS Lett. 487 (1), 66-70 (2000)
 MEDLINE 20584721
 PUBMED 11152886
 REFERENCE 3 (bases 1 to 892)
 AUTHORS Genoscope.
 TITLE Direct Submission
 JOURNAL Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage,
 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
 seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
 This GSS is part of a random genomic sequencing program of thirteen
 yeast species: *Saccharomyces bayanus* var. *uvorum*, *Saccharomyces*
exiguus, *Saccharomyces servazzii*, *Zygosaccharomyces rouxii*,
Saccharomyces kluyveri, *Kluveromyces thermotolerans*, *Kluveromyces*
lactis var. *lactis*, *Kluveromyces marxianus* var. *marxianus*, *Pichia*
angusta, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*,
Candida tropicalis and *Yarrowia lipolytica*. Genomic inserts of 3 to
 5 kb were prepared and both extremities were sequenced. See
 keywords for description of this sequence and for the sequence of
 the other extremity of this insert.

FEATURES
 source
 1. 892
 /organism="Kluveromyces lactis"
 /strain="CLIB 210"
 /variety="lactis"
 /db_xref="taxon:28985"
 /clone="BA0AB037D05"
 /clone_1id="BA0AB"
 /clone_2id="BA0AB"

BASE COUNT 274 a 275 c 120 g 222 t 1 others
 ORIGIN

Alignment Scores:
 Pred. No.: 0.00854 Length: 892
 Score: 131.00 Matches: 59
 Percent Similarity: 36.52% Conservative: 48
 Best Local Similarity: 20.14% Mismatches: 160
 Query Match: 5.53% Indels: 26
 DB: 17 Gaps: 6

US-09-889-314-2 (1-496) x CNS075P4 (1-892)

Oy 8 SerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeuThrSerThrPro 27
 |||||
 Db 45 TCCAGCACTGATTCCTTCAGAACAACTTCAGCACTTCCTCCGAGAACAACTTCAGCACT 104
 Oy 28 GlnGlyValProGlnGlnAspLysLeuSerGlnGlnThrLysGlnIleGlnIleThr 47
 |||||
 Db 105 GATTCCTTCGAAACAACCTTCGAGCACTTCCTCCGAGAACAACTTCAGCACTTCCTCT 164
 Oy 48 ArgGlnGlyLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGlyValAserGlyLys 67
 |||||
 Db 165 GAACACCTTCAGCACTTCCTTCGAGAACAACTTCAGCACTTCCTTCGAGAACAACTTC 224
 Oy 68 AspLysThrSerSer-----ThrLysThrGlnThrAlaProGlnGlnGlyValAla 85
 |||||
 Db 225 TCAACCACTTCCTTCGAGAACAACTTCAGCACTTCCTTCGAGAACAACTTCAGCACT 284
 Oy 86 AlaGlyLysGlnSerSerGlnSerGlnLysAlaGlyValAspThrGlyValSerGlyAla 105
 |||||
 Db 285 TCTTCGAGAACAACTTCAGCACTTCCTTCGAGAACAACTTCAGCACTTCCTTCGAG 344
 Oy 106 AlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlnGlu 125
 |||||
 Db 345 GAAACTTCAGCACTTCCTTCGAGAACAACTTCAGCACTTCCTTCGAGAACAACTTC 404
 Oy 126 AlaSerLysSerMetGlu-----SerThrLeuGlnSerLeuSerLeuSerAlaAla 143
 |||||
 Db 405 TCAACCACTTCCTTCGAGAACAACTTCAGCACTTCCTTCGAGAACAACTTCAGCACT 464
 Oy 144 GlnMetLysGlnValGlnAlaValAlaValAlaAlaLeuSerGlyLysSerSerGlySer 163
 |||||
 Db 465 TCTTCGAGAACAACTTCAGCACTTCCTTCGAGAACAACTTCAGCACTTCCTTCGAG 500

Oy 164 AlaLysLeuGlnThrProGlnLeuProLysProGlnValThrProArgSerGlnValIle 183
 |||||
 Db 501 CAGCCTTCGATACCGACTTCCTTCAGAACAACTTCAGCACTTCCTTCGAGAACAACTTC 557
 Oy 184 GlnIleGlyLeuAlaLeuAlaLeuAlaIleGlnThrLeuGlnGlyValAlaThrLysSerAla 203
 |||||
 Db 558 -----TCAACAACCTGATTCCTTCGAGAACAACTTCAGCACTTCAGCACT 593
 Oy 204 LeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnLeu 223
 |||||
 Db 594 TCTTCGAGAACAACTTCAGCACTTCCTTCGAGAACAACTTCGAGAACAACTTCAGCACT 653
 Oy 224 LysGlnAlaIleLysIleAspLysGlnArgGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 243
 |||||
 Db 654 GAACACCTTCAGCACTTCATCA---GAACAGCCTTCGAGAACAACTTCAGCACTTCAGCA 710
 Oy 244 Gln-----LysSerLysAspLeuGlnGlnGlnGlnGlnGlnGlnGlnGlnGlnGln 259
 |||||
 Db 711 CAACCTTCAGCACTTCATCAAGAACAGCCTTCAGCACTTCAGCACTTCAGCACTTCAGCA 770
 Oy 260 MetIleAlaValSerValAlaIleThrValIleSerIleValAlaAlaIlePheThrCys 279
 |||||
 Db 771 ACTTATTCGTTGACTACCGCTTACCACTGTAATGCTGTGACCAATTTACAGCACA 830
 Oy 280 GlyAlaGlyLeuAlaGlyLeuAlaGlyLeuAlaAlaVal 292
 |||||
 Db 831 ACTGTCCATTTGAATCCAGTACGTAAGCGCTGTC 869

RESULT 24
 CNS05HLZ 1101 bp DNA linear GSS 26-MAY-2000
 LOCUS CNS05HLZ
 DEFINITION Tetradodon nigroviridis genome survey sequence T3 end of clone
 010E07 of library A from Tetradodon nigroviridis, genomic survey
 sequence.
 ACCESSION AL337760.1 GI:8231518
 VERSION AL337760.1
 KEYWORDS GSS: genome survey sequence.
 SOURCE Tetradodon nigroviridis.
 ORGANISM Tetradodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorphae; Acanthopterygii; Percomorphae; Tetraodontiformes;
 Tetraodontidae; Tetraodon.
 1 (bases 1 to 1101)
 Roest-Crollius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fitzames,C., Wincker,P., Brottier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.
 Human gene number estimate provided by genome wide analysis using
 Tetradodon nigroviridis DNA sequence
 Unpublished
 2 (bases 1 to 1101)
 Roest-Crollius,H., Jallion,O., Dasilva,C., Fitzames,C., Fisher,C.,
 Bouneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
 Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetradodon nigroviridis
 Unpublished
 3 (bases 1 to 1101)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000)
 This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetradodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.
 location/Qualifiers
 1. 1101
 /organism="Tetradodon nigroviridis"
 /db_xref="taxon:99883"
 /clone="010E07"
 /clone_1id="A"
 /note="Genoscope sequence ID : C0NA010AC04A1-end : T3"

FEATURES
 source

DEFINITION T7 end of clone BB0AA002G01 of library BB0AA from strain CBS 4732

ACCESSION AL431231

VERSION AL431231.1 GI:12214643

KEYWORDS GSS.

SOURCE *Pichia angusta*.

ORGANISM *Pichia angusta*

REFERENCE 1 (bases 1 to 888)
Soucier,J.L., Aigle,M., Artiguenave,F., Blandin,G., Boloite-Fukuhara,M., Bon,E., Brothier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B., Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S., Saurin,M., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,P. and Weissenbach,J.

AUTHORS

TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies

JOURNAL FEBS Lett. 487 (1), 3-12 (2000)

MEDLINE 20584711

PUBMED 11152876

REFERENCE 2 (bases 1 to 888)
Blandin,G., Llorente,B., Malpertuy,A., Wincker,P., Artiguenave,F. and Dujon,B.

AUTHORS

TITLE Genomic exploration of the hemiascomycetous yeasts: 13. *Pichia angusta*

JOURNAL FEBS Lett. 487 (1), 76-81 (2000)

MEDLINE 20584723

PUBMED 11152888

REFERENCE 3 (bases 1 to 888)
Direct Submission

AUTHORS

TITLE Submitted (08-SEP-2000) Genoscope - Centre National de Séquençage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : sequef@genoscope.cns.fr - Web : www.genoscope.cns.fr)

JOURNAL

COMMENT This GSS is part of a random genomic sequencing program of thirteen yeast species: *Saccharomyces bayanus* var. *uvaurum*, *Saccharomyces exiguus*, *Saccharomyces servazzii*, *Zygosaccharomyces rouxii*, *Saccharomyces kluyveri*, *Kluyveromyces thermotolerans*, *Kluyveromyces lactis* var. *lactis*, *Kluyveromyces marxianus* var. *marxianus*, *Pichia angusta*, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*, *Candida tropicalis* and *Yarrowia lipolytica*. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.

FEATURES

source location/Qualifiers

1..888

BASE COUNT 225 a 190 c 273 g 194 t 6 others

ORIGIN

Alignment Scores:

Pred. No.: 0.00674 Length: 888

Score: 132.00 Matches: 58

Percent Similarity: 38.82% Conservative: 41

Best Local Similarity: 22.75% Mismatches: 122

Query Match: 5.57% Indels: 34

DB: 17 Gaps: 6

US-09-889-314-2 (1-496) x CNS076CP (1-888)

QY 5 SerIleSerSerSerGergIProaSPaNGlnLysAsnIleMetSerGlnValLeuThr 24

DB 813 TCTGCACAAAGCTCATTTTGGCGAAGCGCTTTTCAGATGCTGGGAACACCTTGTCCACAGGCC 754

QY 25 SerThrProGlnGlyValProGlnGlnAspLysLeuSerGlyAsnGlnIle 44

DB 753 TCCCGTAGCGGGGAGATATCTCAATCAGCCAGCAGACATGCTCCACTGCTTAGATAT 694

QY 45 -----GlnGlnThrArgGlnGlnGlySerThrGluMetGluSerAspAlaThrIleAla 62

DB 693 AATGCCAGCTCTCAGTCGATCTATCAGTACGAGGGCATCCAGCTTGGCATCATCTTC 634

QY 63 GlyAlaSerGlyLysAspLysThrSerThrThrLysThrGlnThrAlaProGlnGln 82

DB 633 GGCTCCAGCAGATGCTCTCTCAGCTTACGATATTCGACAGCAGCAGCGCTTCACCTAT 574

QY 83 GlyValAlaIleAlaGlySerGluSerGlu-----SerGlnLys 95

DB 573 GGCTTACGCTCTTTCAATATCATCATCAGATACATTCAGAGGGCTTTACGCCAAT 514

QY 96 AlaGlyAlaSerThrGlyValSer-----GlyAlaAlaIleThrThrAlaSerAsn 112

DB 513 GCGGAGCGCTCAAGTGCATTTCCAGCCAAACAGACATCCAGCACTTGTTCGAC 454

QY 113 ThrAlaThrLysIleAlaMetGlnThrSerIleGlnGlnAlaSerLysSerMetGluSer 132

DB 453 GGTTCAGAGCTCAATGTGCTCTCTACC-----GCTTCTGGCACTGTCCCTAGT 406

QY 133 ThrLeuGluSerLeuGlnSerLeuSerAlaAlaGln----- 144

DB 405 GGAGCTAGTGGAAATTCCTCACTGTTCTCGCGCAGCTCGAGTAATTTTCAGTGTGCC 346

QY 145 -----MetLysGluValGluAlaValAlaValAlaIleAlaLeuSerGlyLysSer 160

DB 345 CCCAGCTCCATTTTCAACAGCGCTTCACGCCAAGTCCAGCGGTGACTTATCTAGTCT 286

QY 161 SerGlySerAlaLysLeuGlnThrProGluLeuProLysProGlyValThrProArgSer 180

DB 285 GCCTTAGCGCTTCACAGTGCACATTCCTCAAGCTGCTTCGACGCGCTTCTTC 226

QY 181 GluValIleGluIleGlyLeuAlaLeuAlaIleGlnThr----- 195

DB 225 AGCTCCCTCAGCGCAACGTCGCCGCGAGAGCTTCCAGCAGCAGCAACAGGGTCT 166

QY 196 ---LeuGlyGlnAlaThrLysSerAlaLeuSerAsnThrIleAlaIleAla 214

DB 165 ATTGTCGAGCAACACTGCTTCCAGTGTCTTACGATATGTTCCAGCCAGGCGACGCTC 106

QY 215 AspGlnThrAsnLysLeuGlyLeuGlyGlnAlaIleLysIle 229

DB 105 ACAATCTAGCTCAGGTGCTCTAGCAAAATACAGATTATCTCTTG 61

RESULT 23

CNS075P4 892 bp DNA linear GSS 07-JUL-2001

LOCUS clone BA0AB037D05 of library BA0AB from strain CLIB 210 of

DEFINITION *Kluyveromyces lactis*, genomic survey sequence.

ACCESSION AL430382

VERSION AL430382.1 GI:12213576

KEYWORDS GSS.

SOURCE *Kluyveromyces lactis*.

ORGANISM *Kluyveromyces lactis*

REFERENCE 1 (bases 1 to 892)
Soucier,J.L., Aigle,M., Artiguenave,F., Blandin,G., Boloite-Fukuhara,M., Bon,E., Brothier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B., Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S., Saurin,M., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,P. and Weissenbach,J.

AUTHORS

TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies

JOURNAL FEBS Lett. 487 (1), 3-12 (2000)

MEDLINE 20584711

PUBMED 11152876

REFERENCE 2 (bases 1 to 892)
Boloite-Fukuhara,M., Lemaire,M., Marmisse,R., Montchoer,R., Duchateau-Nguyen,G., Lemaire,M., Marmisse,R., Montchoer,R., Robert,C., Termier,M., Wincker,P. and Wesolowski-Louvel,M.

QY	286	LeuAlaAlaGlyAlaAlaValGly----	AlaAlaAlaAlaGlyGlyAlaAlaGlyAlaAla	304
	:::	:::		
Db	512	GTTCGGCAGAGAGCTCGGTGGGTAGAGACTCGCAGTAGAGAGTCTCGGTAGAGTCCGA	453	
QY	305	AlaAlaThrThrValAlaAlaThrGlnIleThrValGlnAlaValAlaGln-----	Ala	321
	:::	:::		
Db	452	GTAGAGTCTCGGTAGAGAGTCTCAGTAGAGTCTCGATGGAGTTTCGATGGAGTCTTA	393	
QY	322	VallySGlnAlaValIleThrAlaAlaValArgGlnAlaIleThrAlaAlaIleLysAlaAla	341	
		:::		
Db	392	GTAGAGCTCTGGGTGCTCGGTATCGGACCATCCAAACACAGAGTGAACACGACCG	333	
QY	342	VallySerGlyIleLysAlaAlaPheIleLysThrLeuValLysAlaIleLysAlaIle	361	
	:::	:::	:::	
Db	332	GAGAGATGGGAGTCTCAGAGGGCTTAGAAGTTGGGTGGCTCAACAGCTCGAGAGTG	273	
QY	362	SerlySGlyIleSerLysValPheAlaLysGlyThrGlnMetIleAlaLysAsnPhero	381	
	:::		:::	
Db	272	--GAGGTGGTCTCTCAGTAGAGAGTGGGACMACAGAAAGAACCACTCGTTGAG	216	
QY	382	LysIleuSerLysValIleSerSerLeuThrSer-----	392	
	:::	:::		
Db	215	GAGTCTCGCTGGAGAGTCTCGCTCGGTGTCGGTGAAGAAGAGACTGAGTACTGA	156	
QY	393	-----LysTrp-----	ValThrValGlyValGlyValVal	403
Db	155	GACATGCCAGACTGGTGTAGACACCGGAGGACAGACCACAGAAAGCCGGGAGTAGTACTG	96	
QY	404	AlaAlaProAlaLeuGly	409	
	:::	:::		
Db	95	TCAGCAACAGAGTGGGA	78	
RESULT 21				
LOCUS	AY107616	1027 bp	mRNA	linear
DEFINITION	Zea mays PC0064762 mRNA sequence.			
ACCESSION	AY107616			
VERSION	AY107616.1	GI:21210694		
KEYWORDS	HTC.			
SOURCE	Zea mays.			
ORGANISM	Zea mays.			
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC clade; Panicoideae; Andropogoneae; Zea.			
AUTHORS	Hainey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V.			
TITLE	Maize Mapping Project/Dupont Consensus Sequences for Design of Overgo Probes			
JOURNAL	Unpublished (2002)			
REFERENCE	2 (bases 1 to 1027)			
AUTHORS	Coe,E.C.			
TITLE	Direct Submission			
JOURNAL	Submitted (25-Apr-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA			
FEATURES	Location/Qualifiers			
source	1..1027			
	/organism="Zea mays"			
	/db_xref="MaizeDB:63478"			
	/db_xref="taxon:457"			
	/clone="PC0064762"			
	/clone_idb="Maize Mapping Project/Dupont Consensus Library"			
	/note="this sequence is part of a project of EST assemblies resulting from the application of public contigs to seed Dupont contigs; this resource was assembled by Dupont as part of a collaboration for the overgo addressing of BACs in conjunction with the Maize Mapping Project"			
BASE COUNT	228 a	324 c	315 g	160 t
ORIGIN				

Alignment Scores:	0.00738	Length:	1027
Pred. No.:	132.50	Matches:	78
Score:	39.80%	Conservative:	41
Percent Similarity:	26.09%	Mismatches:	117
Best Local Similarity:	5.59%	Indels:	64
Query Match:	11	Gaps:	12
DB:			
US-09-889-314-2 (1-496) x AY107616 (1-1027)			
QY 24 ThrSerThrProGlnGlyValProGlnGlnAspLeuSerGlyAsnGluThrLysGln 43			
DB 54 ACAGAAACACCAAG-----CCCCAAGGCCCAACTCACCACTCCCAATGAGAAAG 107			
QY 44 IleGlnIthrArgGlnGlyLysAnthrGluMetGluSerAspAlaThrIleAlaGly 63			
DB 108 GC-GACGAACCCCTGCCCTCCGCCAGCCGCCGCCGAGACGAGACGGCTACACGGCGC 166			
QY 64 AlaSerGlyLysAspLysThrSerSerThrThrLysThrGln 78			
DB 167 GTGCCGAGAGAGACAGACAGTCCCATCCACCCCGCAAGTCTACGTACGCTGCC 226			
QY 79 -----AlaProGlnGlnIleValAlaIleLysGlnSerGln 94			
DB 227 GACGCAACCCCGCTCTCTACAGCGCGCTCGTGAAGAGAGAG----- 271			
QY 95 LysAlaGlyAlaAspThrGlyValSerGlyAlaAlaAlaThrThrAlaSerAnThrAla 114			
DB 272 -----GGCGGTGCCACACCGCAAGTCTACGCA 304			
QY 115 ThrLysIleAlaMetGlnThrSerIleGluGlnAlaSerLysSerMetGluSerThrLeu 134			
DB 305 GCT---GTGCCGCCCAAGCCGATCGAGAC----- 334			
QY 135 GluSerLeuGlnSerLeuSerAlaIleGlnMetLysGluValGluAlaValAla 154			
DB 335 -----CTCGCGCCGCCCAAGCTCGACCTCGAGAGAACCTCGTAGGCC 379			
QY 155 AlaLeuSerGlyLysSerSerGlySerAlaLysLeuGlnThrProGluLeuProLysPro 174			
DB 380 CCGCGGAGAACAACTCCTCGCCGAGAG----- 409			
QY 175 GlyValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGln 194			
DB 410 -----ACGACCGCAGCGAGGAGATC-----TTACGACGAGCCGGAGAGGCTC 454			
QY 195 ThrLeuGlyGlu---AlaThrLysSerAlaLeuSerAnThrAlaSerThrGlnAlaGln 213			
DB 455 ACCATCGCGAGTGGCCGCCACCTCTGTGCCGAGAGAGGTGCTCTCGCCGCGAG 514			
QY 214 AlaAspGluThrAsnLys---LeuGlyLeuGlnLysGlnAlaIleLysIleAspLysGln 232			
DB 515 GTCGAACGCTCGATGCTTCTCGAGGTCAGGAAGGGGAGACGACCTGGACACAG 574			
QY 233 ArgGluGlu-----TyrGlnGluMetLysAlaIleGluGlnLysSerLysAspLeuGln 250			
DB 575 AGCAGCAGAAAGTCCGCCAAGAGGTGATGCTGTGCCGACGACGACGACCTCAAGCTAGAG 634			
QY 251 GlyThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIle 270			
DB 635 -----GAGAGATTCAGAGCCCTCGAAGGCTTCGCCGCCGCCACCAACAG 685			
QY 271 SerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAla 290			
DB 686 GAGAGGGAAGCTCGCCG---GCTTCTAGAGGCCCCCAAGGAAGGGAGAGTGGCTGGGCTG 742			
QY 291 AlaValGlyAlaAlaIleAlaGlyAlaAlaGlyAlaAlaGlyAlaAlaAlaAlaThrThrVal 309			
DB 743 GGGATGGCGGTGCTGGCTGGCGCGCGCGGACGCGCTGCATCATCTAGTT 799			
RESULT 22			
LOCUS	CNS076CP/c	888 bp	DNA linear GSS 07-JUL-2001

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 53 Row: 9 Column: 10
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis

This clone has the following problem: incomplete processing.

FEATURES
 Location/Qualifiers

1..2589
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:5364153"
 /issue_type="Eye, retina, mouse strain C57BL/6"
 /clone_lib="NID_MGC_94"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6"
 BASE COUNT 645 a 703 c 615 g 626 t
 ORIGIN

Alignment Scores:

Pred. No.: 0.0154 Length: 2589
 Score: 135.00 Matches: 93
 Percent Similarity: 32.36% Conservative: 40
 Best Local Similarity: 22.63% Mismatches: 155
 Query Match: 5.70% Indels: 123
 DB: 11 Gaps: 15

US-09-889-314-2 (1-496) x BC028951 (1-2589)

QY 76 ThrGluThrAlaProGlnGlnGly---ValAlaAlaGlyLysGlnSerSerGln 94
 Db 41 AGTTCACAGACTCCAGAGCCCAATCTCTGCACAGATCAAAAGCTGTGAGCAAGCT 100
 QY 95 LysAlaGlyAla-----AspThrGlyValSerGlyAlaAlaAlaThrTrpAlaSer 111
 Db 101 AGGACCGGAAACATCATCAATGAGCAAAACAGAAAGATGCTGCGCCACAGACCCCC 160
 QY 112 AsnThrAlaThrLysIleAlaMetGlnThrSerIleGlnGlnAlaSerLysSerMetGlu 131
 Db 161 ACCACCGCAGGCCCA---GCCAGGTCCCGCCACAGCAACCGCAAGTGAAGTACAGAC 217
 QY 132 SerThrLeuGlnSerLeu-GlnSerLeuSerAlaAlaGlnMetLysGluValAlaAla 151
 Db 218 TCCACAGCCCCACAGCAACAGCAG--- 242
 QY 151 ValValAlaAlaLeuSerGlyLysSerSerGlySerAlaLysLeuGlnThrProGlu 171
 Db 243 -----AGTCTCAGCTTACACAGCTCAGC 268
 QY 171 uProLysProGlyValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAla 191
 Db 269 CCCACGGCTGGGGCTTTGCTCACAGGCACCAAGT6--ACCAACCTTCAGGTGGCCG 325
 QY 191 AlaIleGlnThrLeuGlyGluAlaThrLysSerAlaLeuSerAsnTyrAlaSerThrG1 211
 Db 326 G-----CTTACCGGAGTCCCTACTCTCTCAAGCTCAGCAGCAAGGAGAGATCA 373
 QY 211 nAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGlnValGlnAlaIleLysIleAsp 231
 Db 374 AACCTCAGACGCCCGCTGCTCAGGTGGCTGGCAAAACCTCCCGTG----- 422
 QY 231 sGluArgGlnGluTyrGlnGlnMetLysAlaAlaGlnGlnLysSerLysAspLeuGlu 251
 Db 422 ----- 422
 QY 251 yThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSe 271
 Db 423 -----GTATCTGTCCACGACCTGTGGTCTC 448
 QY 271 rIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAla 291
 Db 449 C-----TCCGCCAGGGGTCCACAGCCCTGCTGTAACGTTGC 484
 QY 291 aValGlyAlaAlaAlaGlyAlaAlaGlyAlaAlaGlyAlaAlaAlaThrThrValAla 311

Db 485 AGGATACAGCGTGGCAFTGGACAGCCCAAAAACAGCAGGCGAGCGGTGGGCCCA 544
 QY 311 rGlnIleThrValGlnAlaValValGlnAlaValLysGlnAlaValIleThrAlaVal 331
 Db 545 ACCGTCATGATCCACAGCTGCTGAGTATTAACACACAG-----ACAGCGCTCA 595
 QY 331 gGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyIleLysAlaPheIle 351
 Db 596 ACACAGC----- 602
 QY 351 sThrLeuValLysAlaIleAlaLysAlaIleSerLysGlyIleSerLysValPheAla 371
 Db 603 -----AAGGCCATCCAGCCACCAAGTTGCCAGGCGCAGGCTGTCTTCAGCAAA 652
 QY 371 s---GlyThrGlnMetIleAlaLysAsnProLysLeuSerLysValIleSerIle 390
 Db 653 GCTGACCACCCAACAGATCACTACTCAGGGGCCGAG---CAGAAAGCTC----- 698
 QY 390 uThrSerLysTrpValThrValGlyValAlaValAlaAlaAlaAlaAlaLeuGly 410
 Db 699 -----GCTTATGCTGCCAGCCAGCCCTTAAAC 727
 QY 410 sGlyIleMetGlnMetGlnLeuSerGlnMetGlnGln-----AsnValAlaG1 426
 Db 728 CCAAGTTCTGACCAACCACTCTCCAGGCCAGCAAACTGCCGCGACTCAGCAGTTCA 787
 QY 426 nPheGlnLysGluValGlyLysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrG1 446
 Db 788 GACTCAGATGACAGGTGCAAACTTCTCAAGTTGTGCACGCAACCAACCTGTGGCCAG 847
 QY 446 nPheTrpGlnAlaLeuSerLysIleAlaSerLysGlnThrGlyLysSerGlnMet 466
 Db 848 TATCAACAAGATTGCATCA-----GCTTCTCAGCAGGCTTCTCCACAGAGGTGACGT 901
 QY 466 rGlnLysAlaThrLysLeuGlyAlaGlnIle 476
 Db 902 CACACAGCAACGCGACAGCGCAACAGGTA 932

RESULT 18
 AA930444 485 bp mRNA linear EST 23-APR-1998
 LOCUS VS59603.r1 Strizagene mouse skin (4937313) Mus musculus cDNA clone
 DEFINITION IMAGE:1150612.5' similar to TR:008884 008884 HIGH-GLYCINE TYROSINE
 KERATIN TYPE II.4.; mRNA sequence.

ACCESSION
 AA930444
 VERSION
 AA930444.1 GI:3079225
 KEYWORDS
 SOURCE
 ORGANISM

Mus musculus
 house mouse.
 EST.
 Mus musculus
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 485)
 REFERENCE
 Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
 Geisler,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
 Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
 Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
 Waterston,R.
 The WashU-HMI Mouse EST Project
 Unpublished (1996)

TITLE
 JOURNAL
 COMMENT

WashU-HMI Mouse EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: mouseest@wustl.edu
 This clone is available royalty-free through LLNL; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 MGI:623820
 Possible reversed clone: similarity on wrong strand
 Seg primer: -28m13 rev1 ET from Amersham
 High quality sequence stop: 216.

```

Db 1502 GACCTAGCCTTT-----GACCTATCCATTCACCCAAACCTGAC 1540
Oy 78 ThrAlaProGlnGlnGlyValAlaAlaGlyLysLysSerSerGlnLysAlaGly 97
Db 1541 TTGGGCCATCCCTGGGA---AACTGTGTGAACCTCTCCCAAGACATCCAAATCACT 1597
Oy 98 AlaAspThrGlyValSerGlyAlaAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIle 117
Db 1598 CCATCTGCTTAATACATCTTTAGTAAATCCACTACTGACATGATGTCAGAAATTCAG 1657
Oy 118 AlawetGlnThrSerIleGlnGlnAlaSerLysSerMetGlnSerThr---LeuGlnSer 136
Db 1658 CAACCTGCAAGAGATTTTGCAGATCTACAGAAAGATTAGAGACTCTGAGACAGAGAGA 1717
Oy 137 LeuGlnSerLysSerAlaAlaGlnMetLysGlnValGlnAlaValAlaVal---Ala 154
Db 1718 AAACAGCTACAGGTGCAACCTCCGAGAGGCGAAGCTGTATGCTTAACACACACT 1777
Oy 155 AlaLeuSerGlyLysSerSerGly---SerAlaLysLeuGlnThrProGlnLeuProLys 173
Db 1778 GAGATTTACAGAAACAGCTCTGACCTCAGCCAGAACTT-----AAA 1819
Oy 174 ProGlyValThrProArgSerGlnValIleGlnIleGlnAlaLeuAlaLysAlaIle 193
Db 1820 GAACTTAGCAACAATACGAGAGGCTATGAAAGAGATCTTAGTGTACAGAGACAGATG 1879
Oy 194 GlnThrLeuGlyGlnAlaThrLysSerAlaLeuSerAsnThr--- 207
Db 1880 AAA---CTCGGTCTTGTCTCAGCTGAAAGCATGATTAATTCATCATTTCCACGAGCTG 1936
Oy 208 -----AlaSerThrGln 211
Db 1937 AGGTCACGAGAAGAGAAATTAATGCTGTAAGCAGATCTGCAGAAATGCAATTACAAAG 1996
Oy 212 AlaGlnAlaAspGlnThrAsnLysLeuGlnLysGlnAlaIleLysIleAspLys 231
Db 1997 AGTGAAGAAATTAAGAGAAAGTGAAGATGAGAGAAACCTGCTAGAGAGAGAGAAA 2056
Oy 232 GlnArg-----GlnGlnLysGlnGlnLysAlaAla----- 242
Db 2057 GGTACAGTGAATTAAGCCACCTGTGAGAGATGACAGAAATGAATTCATTTGCTCT 2116
Oy 242 ----- 242
Db 2117 GTTATTGAGAAATATGCAATTAAGGAGAAACATTTTGTGAGAAATACCAAGAACCCAA 2176
Oy 243 -----GlnGlnLysSerLys 247
Db 2177 GAAGAATCATGAAATTAAGACACACTTAAAGTCAAGATGACACAGAGAACCCAGTGAT 2236
Oy 248 AspleuGlnGlyThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIle 267
Db 2237 GAAGTGTGAGACATGAAGAACCCATGATGATGATGATGATGATGATGATGATGATG 2296
Oy 268 ThrValIleSer----- 271
Db 2297 AGCGAGCTCTCAGCTGTACAAAGAACCCAGCTGAGCTGAGAGATTACAGAAAGAG 2356
Oy 272 -----IleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAla 287
Db 2357 AAATCTTAGAGAGATGTCACAGCTGAATATATCCATTAAGCAGACAGATGAGAACTGATG 2416
Oy 288 AlaGlyAlaAlaValGlyAlaAlaAlaAlaGlyAlaAlaGlyAlaAlaAlaAlaAla 306
Db 2417 CAATTGACAAACGCTGCTCCAGGCTAAAGCAGAGATGCTCTGAAATGAGAGTCTCAG 2476
Oy 307 ---ThrThrValAlaThrGlnIleThr---ValGlnAlaValAlaGlnAlaValLysGln 324
Db 2477 TATTCAAAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2536
Oy 325 -----AlaValIleThrAlaValAlaGlnAlaIleThr 335

```

```

Db 2537 AACTGTCTCTTATCACAGAACATTTTGCAGATGATTAACCAAGCTGCGG-----ACT 2587
Oy 336 AlaAlaIleLysAlaAlaValLys---SerGlyIleLysAlaPheIleLysThrLeuVal 354
Db 2588 GCAGCAAAAGAGATGAGAAAGAAAATTAAGCAATCTTAAGAACACCTTGCAGAGCAAGAA 2647
Oy 355 LysAlaIleAlaLysAlaIleSerLysGlyIleSerLysValPheAlaLysGlyThrGln 374
Db 2648 GTGGAAGTACGAAGAGCTGAGAAACACTCTTAGAAGAGAAAGAGCTGATGATGATGCA 2707
Oy 375 MetIleAlaLysAsnPheProLysLeuSerLysValIleSerSerLeuThrSerLysTyr 394
Db 2708 ATGTGTACCTCGGTCT-----TCTATGAAGAAACTCCAGTCATCTTAGAGAGTGA--- 2758
Oy 395 ValThrValGlyValGlyValAlaValAlaAlaProAlaLeuGlyLysGlyIle----- 412
Db 2759 -----GTCAGTGTGTGTCATCGAATTAAGGAATCTGTAAGAG 2800
Oy 413 -----MetGlnMetGlnLeuSerGlnLysGlnGlnAlaAlaGlnPheGlnLys 429
Db 2801 AAAGAGAGAGCTCATTCAGAGTGTTCAGATTAAGAGTGAAGTGTGTCAGAGTGAAGAA 2860
Oy 430 Gln-----Val 431
Db 2861 GAAAGAGAAATTAATTCAGCTCTTGAATCCAAGAGCAAGAGATGAATGAATCTTCTG 2920
Oy 432 GlyLysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPheThrGlnAla 451
Db 2921 CAAAATTCACAGCAAGCTCAGAGAAAGCTTCCAGAAATGAAGAAATGCGTACAGAGCTCT 2980
Oy 452 SerLysIleAlaSerLysGlnThrGlyGlnSerAsnGlnMetThrGlnLysAlaThrLys 471
Db 2981 TCAAAACTGAGAGAAATGAATTAAGATTAATGATGATGATGATGATGATGATGATGATG 3040
Oy 472 Leu 472
Db 3041 TTG 3043

```

RESULT 17
LOCUS BC028951 2589 bp mRNA linear HTC 07-AUG-2002
DEFINITION Mus musculus, similar to E1A binding protein p400, clone

ACCESSION BC028951
IMAGE:5364153, mRNA.
VERSION BC028951.1 GI:22137754
KEYWORDS HTC.
SOURCE house mouse.
ORGANISM Mus musculus.
REFERENCE 1 (bases 1 to 2589)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (01-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs@remail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amgdbcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found

end enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was constructed by Life Technologies. Contact: Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax: (1) 301 610 8371 Email: fliang@lifetech.com URL: http://fulllength.invitrogen.com"

BASE COUNT 511 a 332 c 274 g 76 t 1 others
ORIGIN

Alignment Scores:

Pred. No.: 0.00408 Length: 1194
Score: 136.00 Matches: 59
Percent Similarity: 38.93% Conservative: 36
Best Local Similarity: 24.18% Mismatches: 115
Query Match: 5.74% Indels: 34
DB: Gaps: 7

US-09-889-314-2 (1-496) x AL534303 (1-1194)

QY 7 SerSerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeuThrSerThr 26
DB 427 TCTTCCACCCCAACCCACACCCCAAC----- 456
QY 27 ProGlnGlyValProGlnGlnAsp-----LysLeuSerGlyAsnGlnLysGlnIle 44
DB 457 ---CAAGGCTTCACAGCAAGCAAGCGGAAAGCAAGCAAGCAAGCAAGCAAGCAAG 513
QY 45 GlnGlnThrArgGlnGlnLysAsnThrGlnMetGlnSerAspAlaThrIleAlaGlyAla 64
DB 514 CAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 573
QY 65 SerGlyLysAspLysThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGlyVal 84
DB 574 AGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 633
QY 85 AlaAlaGlyLysGlnSerSerGlnSerGln-----LysAlaGlyAlaAspThrGlyVal 102
DB 634 CAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 693
QY 103 SerGlyAlaAlaAlaThrThrAlaSerAsnThrAlaThrLysIleAlaMetGlnThrSer 122
DB 694 AGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 747
QY 123 IleGlnGlnAlaSerLysSerMetGlnSerThrLeuGlnSerLeuGlnSerLeuSera 142
DB 748 AAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 807
QY 143 AlaGlnMetLysGlnValGlnAlaValAlaAlaLeuSerGlyLysSerSerGly 162
DB 808 GCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 834
QY 163 SerAlaLysLeuGlnThrProGlnLeuProLysProGlyValThrProArgSerGlnVal 182
DB 835 GCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 894
QY 183 IleGlnIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeuGlnGlyAlaThrLysSer 202
DB 895 AAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 939
QY 203 AlaLeuSerAsnThrAlaSerThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlnLeu 222
DB 940 CAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 999
QY 223 GlnLysGlnAlaIleLys-----IleAspLysGlnArgGlnGlnLysGlnIleMetLysAla 241
DB 1000 AACAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1059
QY 242 AlaGlnGlnLys 245
DB 1060 GCAAGCAAGCAAG 1071

RESULT 16

BC028681 3344 bp mRNA linear HTC 01-MAY-2002
LOCUS Homo sapiens, retinoic acid induced 14, clone IMAGE:4822221, mRNA.
DEFINITION BC028681
ACCESSION BC028681.1 GI:20379528
VERSION
KEYWORDS
SOURCE
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
1 (bases 1 to 3344)
Strausberg, R.
Direct Submission
Submitted (29-APR-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REFERENCE

JOURNAL
AUTHORS
TITLE
COMMENT
REMARK
NIH-MGC Project URL: http://mgc.ncl.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
CDNA Library Preparation: Michael J. Brownstein (NHGRI) & Shiraki Toshitsuki and Piero Carninci (RIKEN)
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNU)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
Contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whitting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNU at: http://image.llnl.gov
Series: IRMK Plate: 33 Row: C Column: 23
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 13470085
This clone has the following problem: frame shifted.
Location/Qualifiers
1..3344
/organism="Homo sapiens"
/db_xref="LOCUSID:26064"
/db_xref="taxon:9606"
/clone="IMAGE:4822221"
/issue_type="Testis"
/clone_lib="NIH_MGC_97"
/lab_host="DH10B"
/note="Vector: pBluescript"

FEATURES

source
BASE COUNT 1170 a 691 c 782 g 701 t
ORIGIN

Alignment Scores:

Pred. No.: 0.0197 Length: 3344
Score: 135.50 Matches: 113
Percent Similarity: 37.10% Conservative: 110
Best Local Similarity: 18.80% Mismatches: 207
Query Match: 5.72% Indels: 171
DB: Gaps: 23

US-09-889-314-2 (1-496) x BC028681 (1-3344)

QY 8 SerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeuThrSerThrPro 27
DB 1346 AGTACTACAGTGTCTTACCTTATGTGATTAAGTTCTGTA----- 1387
QY 28 GlnGlyValProGlnGlnAspLysLeuSer-----Gly 38
DB 1388 -----GCTGACCAACAGATCTTCTCTCTATTCAGCAAGCAAAAGTCTCCCTTACC 1441
QY 39 AsnGlnThrLysGlnIleGlnGlnThrArgGlnGlnLysAsn-----ThrGlnMetGlnSer 57
DB 1442 TTACACATTAAGAGGTTTACAGATAAATTACAGCCCAATCCCAAGAGGCGGAGCA 1501
QY 58 AspAlaThrIleAlaGlyAlaSerGlyLysAspLysThrSerSerThrThrLysThrGlu 77

Oy	250	GlUGlyThrMetAspThrValAsnThrValMetIleAlaSerValAlaIleThrVal	269
Db	842	-----CTTATGATGCCCGAAAGGTTCGGTG-----	868
Oy	270	IleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaGly	289
Db	869	-----GTCCGCAGTAACGTTGATGTCCGC-----	892
Oy	290	AlaAlaValGIYAlaAlaAlaAlaGlyLysAlaGlyAlaAlaAlaAlaAlaThrVal	309
Db	893	-----AAGGTTGTGCTGTCCCTCAAGCAGGTGGT-----GCCCGTGCATT	937
Oy	310	AlaThrGlnIleThrValAlaGlnAlaValAlaGlnAlaValLysGlnAlaVal--IleThr	328
Db	938	GTGACCAGAGTGCAGCCCATGTCGTCCCTCCAGCGTGTGATGAGAGGCTTCAGATCCTT	997
Oy	329	AlaValAlaArgGlnAlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyIleLysAla	348
Db	998	CCCTTGGAGGAGCGTTGTCTGAACCTCACATCTTCTGTGACACACACTGCAGAACAGAT	1057
Oy	349	PheIleLysThr-----LeuValLysAlaIleAlaValAlaIleSerLysGlyIleSer	366
Db	1058	ATCATCATGTTGTGACCATGATGAGAAGATGAAGACAATGCCATTGTGTGCACATTGGC	1117
Oy	367	LysValPheAlaLysGlyThrGlnMetIleAlaLysAsnProLysLeuSerLysVal	386
Db	1118	CAC TTG C A A T G A A A T T G A T G C T C C G C C T T G A G A C C T C C T G C G G C G C A A G G C A T C	1177
Oy	387	--IleSerSerLeuThrSerLysTrpVal-----ThrValGlyValGlyValValVal	403
Db	1178	A C C A T T G A A C C C C A G A C T G A C C G C T G G G T G T T C C C G A G A C C A C A C T G C A T T G T C	1237
Oy	404	AlaAlaPro-----AlaLeuGlyLysGlyIleMetGlnMet	415
Db	1238	C T T G C T G A G G G T G C G C C T A T G A C C T T G A G N T G C T A C T G G C A C T A G C T T T G T C A T G	1297
Oy	416	GlnLeuSerGluMetGlnGlnAsnValAlaGlnPhe-----GlnLysLysVal	431
Db	1298	T C C T G C T C A T T C A C T A C C A G G C A C A T T G C C C A A C T T G A A C T G T G A A G A G A A G C T C T	1357
Oy	432	GlyLysLeuGlnAlaAlaAlaAspMetIleSerMetPheThrGlnPheTrpGlnAla	451
Db	1358	G G C A A G T A G A A A A A A G S T G A T G T G C C C C C A A G A C A C C T T G A T -----	1402
Oy	452	SerLysIleAlaSerLysGlnThrGlyLysSerAsnGluMetThrGlnLysAlaThrLys	471
Db	1403	G A G A A G G T T G C G C T C C A C T T G G C -----AAG	1432
Oy	472	LeuGlyAlaGlnIleLeuLysAlaTryAlaAla-----IleSerGlyAlaIle	487
Db	1433	C T T G G T G C C A A C T G A C C A A G S T C A C A A G T O T C A G G C G A C T A A T C A G C G T G C G A T C	1492
Oy	488	AlaGlyAlaHisLysThrAsnAsnPhe	496
Db	1493	G A G G T C C C T A C A A G C C T G C C A C T A C	1519
RESULT 12			
LOCUS	BC016144	4272 bp	mRNA linear HTC 24-OCT-2001
DEFINITION	Homo sapiens, clone IMAGE:3921754, mRNA.		
ACCESSION	BC016144		
VERSION	BC016144.1 GI:16359374		
KEYWORDS	HTC.		
SOURCE	Homo sapiens.		
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
REFERENCE	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 4272)		
TITLE	Strausberg,R.		
JOURNAL	Submitted (22-Oct-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,		

```

REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
              Email: gcapbs-remail.nih.gov
              Tissue Procurement: ATCC/DCTD/DTF
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Genome Sequence Centre,
              BC Cancer Agency, Vancouver, BC, Canada
              info@qsc.bc.ca
              Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
              Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
              Letlicia Hsiao, Martin Krzywinski, Reta Kutsoke, Oliver Lee, Soo
              Sen Lee, Victor Ling, Carrie Mathewson, Candace McLeavy, Steven
              Ness, Pawan Pandoh, Anna-Lisa Parbhoo, Parvaneh Saedi, Jacquelin
              Scheidt, Duane Smalius, Michael Smith, Lorraine Spence, Jeff Stottm
              Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
              George Yang, Scott Zuyderduyn, Marco Marra.

FEATURES
Source
    /organism="Homo sapiens"
    /db_xref="taxon:9606"
    /clone="IMAGE:3921754"
    /tissue_type="Skin, melanotic melanoma."
    /clone_lib="NIH_MGC_72"
    /lab_host="DH10B"
    /note="Vector: pCMV-Sport6"

BASE COUNT   1156 a   1221 c   1352 g   543 t

ORIGIN
Alignment Scores:
Pred. No.:       0.00987          Length:        4272
Score:           140.00          Matches:         118
Percent Similarity: 33.12%       Conservative:     89
Best Local Similarity: 18.88%    Mismatches:     257
Query Match:      5.91%          Indels:          162
DB:               11            Gaps:             19

US-09-889-314-2 (1-496) x BC016144 (1-4272)
QY      7 serSerSerSerGlyProAspAsnGlnLysAsn-IleMetSerGlnValLeuThrSerTh
      :::: ||||| |::: |||||
Db      614 GCCGACAGCCTCCAGCGAGGAGCACACCCTCCAGTCGACGTGTGAGACGACGTGGAGGGGAAA 673
QY      26 rProGlnGlyAlaProGlnGlnAspLysLeuSerGlyAsnGlnLurThrLysGln----- 43
      ||||| | || ||||| ||| |||||
Db      674 CCTCAAGTA-AAACCAGCCCAGGTAAAGCCTCATCACTTCTACTAAGAAGTCTCCAGC 732
QY      43 ----- 43
Db      733 AAGAAAGGCCGCCAGCCCTCGGAAGGTGGGGATGTGACACCCAGTCCAAAGAGG 792
QY      44 ----IleGlnGlnThrArgGlnGlyLysAsnThrcLmetGluSerAspAlaThrIleAl 62
      :::: :::: ||||| ||||| |||||
Db      793 GGCCCTGCCCCAGCCAAGAGGGCCAAAGAACCCAGAGAGAGAGTGACAGAGTAGTAGAGA 852
QY      62 agLYlaSerGlyLysAspLysThrSerSerThrThrLysThrcLunThrAlaProGlnGl 82
      ||||| :::: ||||| :::: |||||
Db      853 GGGATCTGAAGTAGAGAGAGAGAGCCCCCTCGAGGAGACAGAAAGCCAGTAAGAGCCTCGA 912
QY      82 nGLyAlaAlaLaagLyLysGluSerSerGluSerGlnLysAlaactLYlaAspThrgLYVa 102
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      913 AAAAATTCTCCAGTCAGTCGCTCTGACCCCTGCAAG-----GGGAC 957
QY      102 lSerGlyAlaAlaAlaThrTrhAla---SerAsnThrAlaThrLysIleAlaMetGlnTh 121
      ||| ||||| ||| ||| :::: |||||

```

Oy	127	-----SerlysermetglusterrhleuglusterleucInser	139
Db	1902	GCAGAGCCCCACATCACAGGAGCCCAAAACATGAGGGCGGCTCTAAGGCCACAGG	1961
Oy	140	LeuSerAlaalaagImetlysGluaVal-----GluaValaValaValaLaLeu	156
Db	1962	ATTGAAGCCACAGAGTCAGTACCACAGGTGTGGAGAACCAGGGCTTAGAGGCCAGGCTC	2021
Oy	157	SerGlyLysSerSergLySerAlaLysleuGluThrProGluLeuProLysProGlyVal	176
Db	2022	CCGGGGATCTCTGTGTCG------CCCATCTCATCTGTGTGCC	2060
Oy	177	ThrProArgserGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIleGlnThrLeu	196
Db	2061	GCAGAGGCTCTGGAAAAGCTTCACATA-----GAGCTGAGAACCCAGAGATCCGTGCTTG	2114
Oy	197	GlyGluAlaThrLysSerAlaLeuSerAsnTYraLaserrThrglnAlaGlnAlaAspGln	216
Db	2115	GAGAGGCTCTCCGGCAG-----CCGAGCGGGAGAGCAGCTCCGGAGCTG	2159
Oy	217	ThrAsnLysLeuGly--LeuGluLysGlnAlaIleLysIleAspLysGluArgGluGlu	235
Db	2160	GAGCGGCGCTTGGGAGAAATGCGAGAGCCGCGAGCGCAGAGCAGCGGCTGGAGAGCGT	2219
Oy	236	TyrGlnGluMetLys-----	240
Db	2220	GTCCGCGAGGCCGAGGAGCGAGCCGGCGGCGGTGCGGTGACACACAGACTG	2279
Oy	241	-----AlaAlaGlnGlnLysSerLysAspLeuGluGlyThrMetAspThrValAsnThr	258
Db	2280	CGGGCGGCTTGAGACAGAGCCCGGAGAGACCTCCAGACCCGGGACTCCCGCTGGGGAG	2339
Oy	259	ValMetIleAlaValSerValaIleThrValIleSerIleValaIlaIlePheThr	278
Db	2340	CTGAGCGACGCTCGCTCGCTGATGATGAGCTCGGCGCAG--CCGCTGTGTCGGAGAGA	2398
Oy	279	CysGlyAlaGlyLeuAlaGlyLeuAlaIaGlyAlaIaValaGlyAlaIa-----	295
Db	2399	GGAGCGGCGGCGCTCGGCGGAGCTGCGCCAGCGGAGGAGCGCGCTGGAGCAGAG	2458
Oy	296	--AlaAlaGlyGlyAlaAlaGlyAlaIa 304	
Db	2459	CCGGAGCTGGAGGTTCTGCGGAGACACT 2488	
RESULT 11			
LOCUS	AY103583	1845 bp	mRNA linear HTC 25-MAY-2002
DEFINITION	Zea mays PCO113569 mRNA sequence.		
ACCESSION	AY103583		
VERSION	AY103583.1 GI:21206661		
KEYWORDS	HTC.		
SOURCE	Zea mays.		
ORGANISM	Zea mays.		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC Clade; Panicoideae; Andropogoneae; Zea. 1 (bases 1 to 1845) Haneey,C.F., Dolan,M., Miao,G.H., Vogel,J.M., Whitsitt,M.S., Arthur,L.W., Hanafey,M., Morgante,M. and Tingey,S.V. Maize Mapping Project/DuPont Consensus Sequences for Design of Overgo Probes Unpublished (2002) 2 (bases 1 to 1845) Coe,E.C. Direct Submission Submitted (25-APR-2002) Maize Mapping Project, University of Missouri, Columbia, MO 65211, USA Location/Qualifiers 1..1845 /organism="Zea mays" /db_xref="Maizedb:63617" /db_xref="taxon:4577" /clone="PCO113569"		
JOURNAL			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			

BASE COUNT	384 a	538 c	514 g	409 t																									
ORIGIN																													
<p>Alignment Scores:</p> <table><tr><td>Pred. No.:</td><td>0.00301</td><td>Length:</td><td>1845</td></tr><tr><td>Score:</td><td>140.00</td><td>Matches:</td><td>112</td></tr><tr><td>Percent Similarity:</td><td>33.27%</td><td>Conservative:</td><td>64</td></tr><tr><td>Best Local Similarity:</td><td>21.17%</td><td>Mismatches:</td><td>223</td></tr><tr><td>Query Match:</td><td>5.91%</td><td>Indels:</td><td>130</td></tr><tr><td>DB:</td><td>11</td><td>Gaps:</td><td>21</td></tr></table>						Pred. No.:	0.00301	Length:	1845	Score:	140.00	Matches:	112	Percent Similarity:	33.27%	Conservative:	64	Best Local Similarity:	21.17%	Mismatches:	223	Query Match:	5.91%	Indels:	130	DB:	11	Gaps:	21
Pred. No.:	0.00301	Length:	1845																										
Score:	140.00	Matches:	112																										
Percent Similarity:	33.27%	Conservative:	64																										
Best Local Similarity:	21.17%	Mismatches:	223																										
Query Match:	5.91%	Indels:	130																										
DB:	11	Gaps:	21																										

US-09-889-314-2 (1-496) x AY103583 (1-1845)

QY	8	SerserSgIyProApsaInGlnLysAsnIleMetserGlnValLeuThrSerThrPro	27
		
Db	203	AeTTCGGCCCGCTCAAGCCCTTCGCCCGCGCTAGAGATCTGGGCTCTCTCCACTGACCA	262
QY	28	GInGlyValProGInGlnAspLysLeuSerGlyAsnGlnuThrLysGlnIleGInGlnThr	47
		
Db	263	TCCACACCGCGCTCCTCATCTGACAGACCTCCACGGCGCTCGGCGCGGAG	310
QY	48	ArgGInGlyLysAsnThrGluMetGluSerAspAlaThrIleAlaGlyLaserGlyLys	67
		
Db	311	-----TCCGCTGGCGCT	322
QY	68	AspLysThrSerSerThrThrLysThrGluThrAlaPro	80
		
Db	323	CCTGCACATCTTCTCCACGACGAGGACACCGCGCGCGCCATCCGCCGCACTCGCGCG	382
QY	81	-----GInGInGlyValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGlyAla	98
		
Db	383	CCGCTTTCGCCCTCGAAGGGGAGAGACCTCCGAGAGATCTGGTGTGACCGAGCGCTGCC	442
QY	99	AspThrGlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThr---AlaThrLysIle	117
		
Db	443	TGCAGCTGGGGCGAGCGGGCGCGCCCGCCGCACTATCTGTCAGACGAGGGCGGCGCCACGC	502
QY	118	AlaMetGlnThrSerIleGluGlnAlaSerLysSerMetGluSerThr-----LeuGlu	135
		
Db	503	TCCTCATCCACGAGGGGTCAAGCGCCGAGGAGATTCACGAGGAACCGCGCAGATCCC	562
QY	136	SerLeuGlnSerLeuSerAlaAlaGlnMetLysGluValaGlnAlaValaValaAla	155
		
Db	563	GACCCGGAGTCCACGACACCGCTGAGTTCAAAGATCGTCTCACATCATCTCCGAGCGG	622
QY	156	LeuSerGlyLysSerSerGlySerAlaLysLeuGlu-----ThrProGlu	170
		
Db	623	CTCAGAGCTGACCCCAAGAAAGTACCGCAAGATGAAGAGAGAGAGCTTCTGCGCTCTGAG	682
QY	171	LeuProLysProGlyValThrProArgSerGluValIleGluIleGlyLeuAlaLeuAla	190
		
Db	683	GAGACCAACACGCGGTGTCAAGAGCTCTACCAAGATGACGAGAGACCGCGCTCTCTTC	742
QY	191	LysAlaIleGlnThrLeuGlyGlnAlaThrLysSerAlaLeuSerAsn---TyrAlaSer	209
		
Db	743	CCTGCCATTAACTACAGGATTCCTGCTACCAAGACCAAGTTTGACACCTGTATGGTTC	802
QY	210	ThrGlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLeuGluLysGlnAlaIleLysIle	229
		
Db	803	CGCCTCTGCTCTCGAT-----GGTCTG	828
QY	230	AspLysGluArgGluGlnuThrGlnuMetLysAlaAlaGluGlnLysSerLysAspLeu	249
		
Db	827	-----ATGAGGGCGCACTGAC-----	841

Alignment Scores:

Pred. No.:	0.00145	Length:	2295
Score:	144.50	Matches:	71
Percent Similarity:	33.33%	Conservative:	23
Best Local Similarity:	25.18%	Mismatches:	111
Query Match:	6.10%	Indels:	77
	11	Gaps:	8

US-09-889-314-2 (1-496) x BC016304 (1-2295)

```

QY 62 AlacglaaaserglylaspserthrsrserthrrhlysthrhlaaProglu 81
    ||| ||||| ||||| |||
Db 1242 GCCCAGCTTACAGGAAAGGGGACACAGCTGCTGGGCTGCTGGCTCTGCC 1183
QY 82 GlnGly-----ValAlaAlaGlyLysGluSerSerGluSerGlnLysAlaGly 97
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1182 CAGGGCTGCTGAGGAGCAGCTGCTGGGCTGCTGCTGCTGAGGACCTGCTGG 1123
QY 98 -----AlaaspthrGly 101
    ||| |||||
Db 1122 ACTGCTGCTGCTGGGACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1063
QY 102 ValserGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 121
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1062 GCTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1003
QY 122 SerIleGluGluAlaSerLysSerMetLysSerThrLysGluSerLysGluSer 141
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1002 GCTACTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 943
QY 142 AlaAlaGlnMetLysGluValGlnAlaVal-----ValValAlaAla 155
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 942 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 883
QY 156 LeuSerGlyLysSerSerGlySerAlaLysLysGluThrProGluLeuProLys 175
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 882 GCTGCTGAGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 844
QY 176 ValThrProAlaSerGluValIleGluIleGlyLeuAlaLeuAlaLysAlaIle 195
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 843 -----GGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 814
QY 196 LeuGlyGlnAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 215
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 813 GCTGGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 757
QY 216 GlnThrAsnLysLeuGlyLeuGlnLysGlnAlaAlaIleLysIleAspLysGlu 235
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 757 ----- 757
QY 236 TyrGlnGlnMetLysAlaAlaGlnLysSerLysAspLeuGlnGlyThrMetAsp 255
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 756 -----GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
QY 256 ValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIleVal 275
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 720 ---GGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 664
QY 276 IlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaGly-----Ala 293
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 663 GTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 604
QY 294 AlaAlaAlaAla-----AlaGlyLysAlaAlaGlyAlaAlaAlaAlaAla 311
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 603 GCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 544
QY 312 GlnIle 313
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 543 AGCGTT 538
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
RESULT 9
CNS07CG2

```

LOCUS CNS07CG2 938 bp DNA linear GSS 08-JUL-2001

DEFINITION T7 end of clone BD0A002F07 of library BD0A from strain CBS 94 of Candida tropicalis, genomic survey sequence.

ACCESSION AL439128

VERSION AL439128.1 GI:12222541

KEYWORDS GSS.

SOURCE

ORGANISM Candida tropicalis.

Candida tropicalis

Eukaryota; Fungi; Ascomycota; Saccharomycetales; Saccharomycetes; Saccharomycetales; mitosporic Saccharomycetales; Candida.

REFERENCE

AUTHORS Souciet, U.L., Aigle, M., Artiguenave, F., Blandin, G., Boly, P., Bouteiller, P., Brothier, P., Casaregola, S., de Montigny, J., Dujon, B., Durand, P., Lepoint, A., Llorente, B., de Montigny, J., Dujon, B., Durand, P., Lepoint, A., Llorente, B., Saurin, W., Tekala, F., Toffano-Nicche, C., Wesolowski-Louvel, M., Wincker, P., and Weissenbach, J.

TITLE Genomic exploration of the hemiascomycetous yeasts: 1. A set of yeast species for molecular evolution studies

JOURNAL FEBS Lett. 487 (1), 3-12 (2000)

MEDLINE 20584711

PUBMED 11152876

REFERENCE

AUTHORS Blandin, G., Ozier-Kalogeropoulos, O., Wincker, P., Artiguenave, F., and Dujon, B.

TITLE Genomic exploration of the hemiascomycetous yeasts: 16. Candida tropicalis

JOURNAL FEBS Lett. 487 (1), 91-94 (2000)

MEDLINE 20584726

PUBMED 11152891

REFERENCE

AUTHORS

TITLE

COMMENT

2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : seque@genoscope.cns.fr - Web : <http://www.genoscope.cns.fr>)

This GSS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces kluyveri, Saccharomyces servazii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Kluyveromyces fragilis, Debaryomyces hansenii var. hansenii, Pichia stipitiphila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.

FEATURES

source

location/Qualifiers

1..938

/organism="Candida tropicalis"

/strain="CBS 94"

/db_xref="taxon:5482"

/clone="BD0A002F07"

/clone_11b="BD0A"

/note="end : 17"

BASE COUNT 155 a 246 c 196 g 339 t 2 others

ORIGIN

Alignment Scores:

Pred. No.:	0.000728	Length:	938
Score:	142.00 <td>Matches:</td> <td>65</td>	Matches:	65
Percent Similarity:	42.97% <td>Conservative:</td> <td>48</td>	Conservative:	48
Best Local Similarity:	24.71% <td>Mismatches:</td> <td>120</td>	Mismatches:	120
Query Match:	5.99% <td>Indels:</td> <td>30</td>	Indels:	30
	17	Gaps:	8

US-09-889-314-2 (1-496) x CNS07CG2 (1-938)

```

QY 56 GluserAspAlaThrIleAlaGlyLysAspLysThrSerSerThrThrLys 75
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 154 GATCGGATGCA-----TCTGCCACTTCTGGGTTCATCGCTGATGCTACTGAA 207
QY 76 ThrGluThrAlaProGlnGlnGlyValAlaAlaGlyLysGluSerSerLysGlnLys 95
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.

FEATURES
source
1. 1094
/organism="Pichia angusta"
/strain="CBS 4732"
/db_xref="taxon:4905"
/clone="BB0AA002F11"
/note="end : r3"

BASE COUNT 320 a 245 c 321 g 207 t 1 others
ORIGIN

Alignment Scores:
Pred. No.: 0.00255 Length: 1094
Score: 147.50 Matches: 75
Percent Similarity: 39.19% Conservative: 70
Best Local Similarity: 20.27% Mismatches: 174
Query Match: 6.23% Indels: 51
Gaps: 9

US-09-889-314-2 (1-496) x CNS076CM (1-1094)

QY 8 SerSerSerGlyProAspAsnGlnLysAsnIleMetSerGlnValLeuThrSerThrPro 27
DB 1008 AGCCAAAGCTCAACTTCCAAACAGGCTTCAGGTGATCCAGCTCTACTGCTGCTTCA 949
QY 28 GlnGlyValProGlnGlnAspLysLeu-----SerGlyAsn 39
DB 948 AGTAATGCTCTGACCTGATACCATTCACAGATCTGCGCAAGTAGTAGGATCGGCTCG 889
QY 40 GlnThrLysGlnIleGlnGlnThrArgGlnGlyLysAsnThrGlnMetGlnSerAspAla 59
DB 888 TCCAGCGGTTCTGACTGTCGACAGGATTATGATCTAGCACTGCTATCTCGATTCCGA 829
QY 60 ThrIleAlaGlyLaserGlyLysAspLysThrSerSerThrLysThrLysThrLysAla 79
DB 828 ACTTCAAGCTGCTGCT-----ACCTCAAGCTCAGGTTTATGCTCAAGCGCC 784
QY 80 ProGlnGlnGlyValAlaAlaGlyLysGlnSerSerGlnLysAlaGlyAlaAsp 99
DB 783 GCCACCTCGATTTATGATTTAGGCTTACTTCAAGCGCGTCCAGGCTCTTCCAGC 724
QY 100 ThrGlyValSerGlyAlaAlaAlaThrThrAlaSerAsnThrLysIleAlaMet 119
DB 723 TCTGCCACTCTCGATCAAGCTTTTCAGCTGCTGCTCCGCGCTCCGCTACT 664
QY 120 GlnThrSerIleGlnGlnValLaserLys-----SerMetGlnSerThrLeu 134
DB 663 TCTGCCCTCAAGTTATGATCAACACAGCTTCTGTTAAGCTTCCGCGCTGCTATA 604
QY 135 GlnSerLeuGlnSerLeuSerAlaAlaGlnMetLysGlnValGlnAlaValAlaAla 154
DB 603 TCAGGTTTGAAGCTGCTCAAGTTATGATCAAGGACTACTGCTAGTTTAAAGCTCT 544
QY 155 AlaLeuSerGlyLysSerSerGlySerAlaLysLeuGlnThrProGlnLeuProLysPro 174
DB 543 CGCGCTCTCGTTCAGTTATGCTCAAGC-----514
QY 175 GlyValThrProArgSerGlnValIleGlnIleGlyLeuAlaLeuAlaLysAlaIleGln 194
DB 513 -----TCTACTGCTGCTGCTTATGAGCTCTGCTCAAGTTACGATTAAAGC 469
QY 195 ThrLeuGlyGlnAlaThrLysSerAlaLeuSerAsnTyraLaserThrGlnAlaGlnAla 214
DB 468 ACCGAGGCTGCTGCT-----TCAAGTTATGATCAAGACCGCAAGCTCT 424
QY 215 AspGlnThrAsnLysLeuGlyLeuGlnLys-----GlnAlaIleLysIleAspLysGlnArg 233
DB 423 -----GGTTTGAAGCTTCCGCTGCTCAATATACAGGCTCAACTCTGCC 382
QY 234 GlnGlnLysGlnGlnMetLysAlaAlaGlnLysSerLysAspLeuGlnGly---Thr 252

DB 381 TCAAGTATGCGCTCAAGCACTGCTGCTCTACTTAAAGCTTCCGCGCTGCTTCCACT 322
QY 253 MetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSerIle 272
DB 321 TATGCTCAAGTTCTCTACTGCTGCTGATCAAGCTATGCTCAAGCTCTGCTGCTCAGGC 262
QY 273 ValAlaAlaIlePheThrLysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaVal 292
DB 261 TCCAGCTCTGCTCAAGTTATGATCAAGCACTGCTGCTGCTGCTGCTGCTGCTGCTCA 202
QY 293 GlyAlaAlaAlaAlaGlyGlyAlaAlaGlyAlaAlaAlaAlaAlaThrValAlaThrGln 312
DB 201 AGTTATGCTCAAGCTCTCTATCTCAGATCAAGCTGCTGCTCAAGCTATGCTCAAGC 142
QY 313 IleThrValGlnAlaValAlaGlnAlaValLysGlnAlaValIleThrAlaValArgGln 332
DB 141 TCTACTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 91
QY 333 AlaIleThrAlaAlaIleLysAlaAlaValLysSerGlyLysLysAlaPheIleLysThr 352
DB 90 -----GGCTCAGGTTTGAAGCTCTGCTCAAGTTATGATCAAGCTTACTGCTGCTG 37
QY 353 LeuValLysAlaIleAlaLysAlaLysSer 362
DB 36 TTTGAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 7
RESULT 8
BC016304/c 2295 bp mRNA linear HTC 05-NOV-2001
LOCUS
DEFINITION Homo sapiens, similar to dentin staphosphoprotein, clone
IMAGE:3844335, mRNA.
ACCESSION BC016304
VERSION BC016304.1 GI:16740886
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
REFERENCE Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (29-OCT-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA.
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: c9abs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Institute for Systems Biology
http://www.systemsbio.org
Contact: amadan@systemsbiology.org
Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: http://image.lnl.gov
Series: IRAL Plate: 25 Row: b Column: 20
This clone has the following problem: frame shifted.
location/Qualifiers
1. 2295
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3844335"
/tissue_type="Placenta, choriocarcinoma"
/clone_id="NIH-MGC-21"
/lab_host="DH10B-R"
/note="Vector: pOTB7"

BASE COUNT 558 a 732 c 560 g 445 t
ORIGIN

Eukaryota: Alveolata; Apicomplexa; Coccidia; Eimeriida; Eimeriidae;
 Eimeria.
 1 (bases 1 to 619)
 Ng,S.T., Jangli,M.S., Shirley,M.W., Tomley,F.M. and Wan,K.L.
 Comparative EST analyses provide insights into gene expression in
 two asexual developmental stages of *Eimeria tenella*
 unpublished (2001)
 JOURNAL
 COMMENT
 Contact: Wan K.L.
 Centre for Gene Analysis and Technology
 Universiti Kebangsaan Malaysia
 43600 UKM Bangi, Selangor DE, Malaysia
 Tel: 6 03 8292997
 Fax: 6 03 8293249
 Email: klawan@krisc.cc.ukm.my
 PCR Primers
 FORWARD: T3
 BACKWARD: T7
 Seq primer: SK.

FEATURES
 source
 Location/Qualifiers
 1..619
 /organism="Eimeria tenella"
 /strain="Houghton"
 /db_xref="taxon:5802"
 /clone="etsHa017"
 /clone_id="ETC11"
 /dev_stage="Sporozoite"
 /lab_host="X11-Blue MRF"
 /note="Vector: Lambda ZAPIT; Site_1: EcoRI; Site_2: XhoI;
 Sporozoites were excysted in vitro from E. tenella H
 oocysts, purified by column chromatography and mRNA
 extracted using a FASTTRACK kit (Invitrogen). cDNA was
 synthesised and a uni-zap xr library was constructed
 using cDNA synthesis kit, ZAP-cDNA synthesis kit and
 ZAP-cDNA Gigapack III Gold cloning kit (Stratagene). The
 library was amplified once through E. coli X11-Blue
 MRF."

BASE COUNT 222 a 163 c 180 g 54 t
 ORIGIN

Alignment Scores:
 Pred. No.: 0.000114 Length: 619
 Score: 147.50 Matches: 69
 Percent Similarity: 36.59% Conservative: 21
 Best Local Similarity: 28.05% Mismatches: 106
 Query Match: 6.23% Indels: 51
 DB: 13 Gaps: 5

US-09-889-314-2 (1-496) x BG929588 (1-619)

252 ThrMetAspThrValAsnThrValMetIleAlaValSerValAlaIleThrValIleSer 271
 111 ::: 111 111 111
 Db 12 ACCAGGGAAGCAGCAAGCTGTACAGCTGCAGCAGCAAGGAGGAGCAGCAGC 71
 272 TleValAlaIleIleThrCysGlyAlaGlyLeuAlaGlyLeuAlaIleGlyAlaIle 291
 111 ::: 111 111 111 111 111
 Db 72 GATGCAGCAGACAGTTCACAGCAACCCAGCAGTTCGAAAGCAGCAGCACTGCAGCG 131
 292 ValGlyAlaIleAlaIleIleGlyAlaIleGlyAlaIle-----AlaAlaIle 306
 111 111 111 111 111 111 111 111
 Db 132 AACGCTGCAGCAAGAACTGCAGCAAAAGCTGCAGCAAAAGCTGTACCAAAAGCTGCAGCA 191
 307 ThrThrValAlaIleThrGlnIleThrValGlnAlaValAlaGlnAlaValLysGlnAlaVal 326
 111 111 111 111 111 111 111 111
 Db 192 AGGACTACAGCAAAAGCTGCAGCAAGAGCTGCAGCAAGAGCTGCAGCAAAAGCTGCAGCA 251
 327 IleThrAlaValArgGlnAlaIleThrAlaAlaIleLysAlaAlaVal-LysSerGlyTil 346
 111 111 111 111 111 111 111 111
 Db 252 AGGACTACAGCAAAAGCTGCAGCAAGAGCTGCAGCAAGAGCTGCAGCAAAAGCTGCAGCA 311
 346 elysAlaIleIleIleThrLeuValLysAlaIleAlaIleAlaIleSerLysGlyIleSe 366
 111 111 111 111 111 111 111 111
 Db 312 AAAAGCTGCAGCAAGAGCTGCAGCAAAAGCTGCAGCAAAAGCTGCAGCAAGAGCTGCAGC 371

Qy 366 rlyValpIleAlaIysGlyThrGlnMetIleAlaIysAsnIleProLysLeuSerLysVa 386
 111 111 111 111
 Db 372 AAAAGCTGCAGCAAAAGCTACGGCAAAAGCTGCAGCAAA----- 411
 Qy 386 lIleSerIleThrSerLysTTPValThrValGlyAlaIleValAlaIleApr 406
 412 -----GCTGCAGCGGCAAGCC 428
 Qy 406 GAlaLeuGlyLysGlyIleMetGlnMetGlnLeuSerGlnMetGln--GlnAsnValAl 425
 111 111 111 111 111 111 111 111
 Db 429 AGCAACCGGAGAGAGATGGTGAATTTCCTGTCAGAGCGCAGCAACATCGCAAAAGCTA 488
 Qy 425 aGlnpHeGlnLysGlyValGlyLysLeuGlnAlaAlaIleAspMetIleSerMetpHeTh 445
 111 111 111 111 111 111 111 111
 Db 489 GCATGTGCAGCAGCAGG----- 504
 Qy 445 rGlnpHeTTPGlnGlnAlaSerLysIleAlaSerLysGlnThrGlyLysSerAsnGlnMe 465
 111 111 111 111 111 111 111 111
 Db 505 -----CAGCAACTGCAGCAGCAGCAGCAGCAGCAGCATCGGAACACCTTCGACCA 554
 Qy 465 tThrGlnLysAlaThrLysLeuGlyAlaGlnIleLeuLysAlaTyrAlaAlaIleSerG 485
 111 111 111 111 111 111 111 111
 Db 555 GCTGCAGCA-GCA-----GCAAGCATATCGGAAGCCGCTGCAGCAGCTGCAGC 601
 Qy 485 yAlaIleAlaGlyAla 490
 111 111 111
 Db 602 ACCAGCAGCAGCAGCA 617

RESULT 7
 CDS076CM/C 1094 bp DNA linear GSS 07-JUL-2001
 LOCUS T3 end of clone BB0AA002F11 of library BB0AA from strain CBS 4732
 DEFINITION of *Pichia angusta*, genomic survey sequence.
 ACCESSION AL431228
 VERSION AL431228.1 GI:12214640
 KEYWORDS GSS.
 SOURCE *Pichia angusta*.
 ORGANISM *Pichia angusta*.
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; *Pichia*.

REFERENCE
 AUTHORS Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
 Bolotin-Fukuhara,M., Bon,E., Broillet,P., Castegnola,S.,
 de-Montigny,J., Dujon,B., Durans,P., Leplingle,A., Llorente,B.,
 Malpertuy,A., Neuvéglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
 Saurin,W., Tekala,F., Toffano-Nicche,C., Wesolowski-Louvel,M.,
 Winkler,P. and Weissenbach,J.
 Genomic exploration of the hemiascomycetous yeasts: 1. A set of
 yeast species for molecular evolution studies
 FEBS Lett. 487 (1), 3-12 (2000)

REFERENCE
 AUTHORS Blandin,G., Llorente,B., Malpertuy,A., Winkler,P., Artiguenave,F.
 and Dujon,B.
 Genomic exploration of the hemiascomycetous yeasts: 13. *Pichia*
angusta
 FEBS Lett. 487 (1), 76-81 (2000)

JOURNAL
 MEDLINE 20584723
 PUBMED 11152888
 3 (bases 1 to 1094)

REFERENCE
 AUTHORS Direct Submission
 TITLE Genoscope.
 JOURNAL Submitted (08-SEP-2000) Genoscope - Centre National de Séquençage,
 2 rue Gaston Crémieux, CP 5706, 91057 Evry cedex, FRANCE. (E-mail :
 seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT
 This GSS is part of a random genomic sequencing program of thirteen
 yeast species: *Saccharomyces bayanus* var. *uvarum*, *Saccharomyces*
exiguus, *Saccharomyces servazii*, *Zygosaccharomyces rouxii*,
Saccharomyces kluyveri, *Kluyveromyces thermotolerans*, *Kluyveromyces*
lactis var. *lactis*, *Kluyveromyces marxianus* var. *marxianus*, *Pichia*
angusta, *Debaryomyces hansenii* var. *hansenii*, *Pichia sorbitophila*,
Candida tropicalis and *Yarrowia lipolytica*. Genomic inserts of 3 to

QY 273 ValAlaAlaIlePheThrcysGlyAlaGlyLeuAlaGlyLeuAlaAlaGlyAlaAlaVal 292
 |||||::: ||| ::||| |||||
 Db 238 GTAGCTGCTGCTAGCTAGCTGCTGCTAGCTGCTAGCTGCTAGCTGCTGCT 297
 QY 293 GLYAlaAlaAlaGlyAlaAlaGlyAlaAlaAlaAlaAlaThrValAlaAlaArg 312
 ||||| ||||| |||||
 Db 298 GTAGCTGCTGCTAGCTGCTGCTGCTAGCTGCTGCTAGCTGCTGCTGCTGCT 354
 QY 313 IleThrValGlnAlaValAlaGlnAlaValAlaValAlaIleThrAlaValArgGln 332
 ||| ||||| ||| |||||::: |||||
 Db 355 ---GTTGTAAGTCTGCTGTTGTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCT 411
 QY 333 AlaIleThrAlaAlaIleAlaAlaAlaValAlaLysSerGlyIleLysAlaPhe---IleLys 351
 :: ||| ::||| |||||::: |||
 Db 412 GTTGTACTGCTGTTGTAAGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 471
 QY 352 ThrLeuValLysAlaIleAlaLysAlaIleSerLys---GlyIleSerLysValPheAla 370
 ::||| |||||::: |||||
 Db 472 GCTGTGTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 531
 QY 371 LysGlyThrGlnMetIle-----AlaLysAsnProLysLeuSerLysValIleSer 388
 ::||| ||| |||
 Db 532 GCTGCTGTAAGTCTGCTGTTGTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 591
 QY 389 SerLeuThrSerLysTrpVal---ThrValGlyValGlyValValAlaAlaAlaProAla 407
 ::||| ::||| ||| ||| |||||
 Db 592 GCTGTACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 651
 RESULT 5
 BM588321 711 bp mRNA linear EST 25-FEB-2002
 LOCUS 17000687322518 A.Gam.ad.cdna.blood1 Anopheles gambiae CDNA clone
 DEFINITION
 1960049697628 5', mRNA sequence.
 ACCESSION
 BM588321
 VERSION
 BM588321.1 GI:18884182
 KEYWORDS
 EST.
 ORGANISM
 African malaria mosquito.
 Anopheles gambiae
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Nematocera; Culicoidae;
 Anopheles.
 REFERENCE
 1 (bases 1 to 711)
 AOTHONS Holt,R.A., Lin,J.-J., Murphy,S.D., Evans,C.A., Kraft,C.L., Charlab
 TITLE 'R', Collins,F.H., Venter,J.C. and Hoffman,S.L.
 JOURNAL Celera Anopheles gambiae EST project
 COMMENT Unpublished (2002)
 Contact: Holt R.A.
 Celera Genomics
 45 W. Gude Dr., Rockville, MD 20850, USA
 Tel: 2404533151
 Fax: 2404534580
 Email: HoltRA@celera.com
 Plate: NU01004AAR row: F column: 14
 Seq primer: M13 Reverse.
 FEATURES
 source location/qualifiers
 1..711
 /organism="Anopheles gambiae"
 /strain="RSP-ST (Reduced susc. to Permethrin - std.
 chromosome)"
 /db_xref="taxon:7165"
 /clone="1960049697628"
 /clone_lib="A.Gam.ad.cdna.blood1"
 /dev_stage="Adult"
 /lab_host="DH10b"
 /note="Vector: pSport1; site_1: salI; site_2: NotI; whole
 adult mosquitoes (mixed sex) frozen on liquid nitrogen 24
 hours after human blood feeding. CDNA inserts >500 bp
 cloned directionally into pSport 1. Not 1 site is 3'.
 Clones available through the Malaria Research and
 Reference Reagent Resource Center (www.malaria.rtrd.org)"
 BASE COUNT 305 a 234 c 135 g 37 t

Alignment Scores:
 Pred. No.: 0.000123 Length: 711
 Score: 148.00 Matches: 64
 Percent Similarity: 38.55% Conservative: 32
 Best Local Similarity: 25.70% Mismatches: 134
 Query Match: 6.25% Indels: 20
 DB: 13 Gaps: 4
 US-09-889-314-2 (1-496) x BM588321 (1-711)
 QY 70 ThrSerSerThrThrLysThrGlnThrAlaProGlnGlnGlyValAlaAlaGlyLysGlu 89
 ||| ||| ||| |||||
 Db 4 ACCGCTCCGGAC 57
 QY 90 SerSerGlnSerGlnLysAlaGlyAlaPheThrGlyValSerGlyAlaAlaAlaThrThr 109
 ||| ||| ||| |||||
 Db 58 -----ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 102
 QY 110 AlaSerAsnThrAlaThrLysIleAlaMetGlnThrSerIleGlnGlnLysSerLysSer 129
 |||::: ||| ||| ||| |||||
 Db 103 GCACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 162
 QY 130 MetGlnSerThrLeuGlnSerLeuGlnSerLeuSerAlaAlaGlnMetLysGluValGlu 149
 ::||| ||| |||
 Db 163 AGAAATGCACTGCTGCTCAACAAAACAGCAGCATCAGTGCA----- 204
 QY 150 AlaValValAlaAlaAlaLeuSerGlyLysSerSerGlySerAlaLysLeuGlnThrPro 169
 ::||| |||||
 Db 205 -----GCAATCACACATGCTCCAGGAAATCAACAGCATCTGCGACATGCACAGCA 258
 QY 170 GlnLeuProLysProGlyValThrProArgSerGluValIleGlnLysLeuAlaLeu 189
 ||||| |||||
 Db 259 TCA-----AATGCACAAATTCACAAATTCACAGGAGCATCAGCAGCAAAATATCGCA 312
 QY 190 AlaLysAlaIleGlnThrLeuGlyGluAlaThrLysSerAlaLeuSerAsnTrpAlaSer 209
 ||| ||| |||
 Db 313 ACACGAC 372
 QY 210 Thr-GlnAlaGlnAlaAspGlnThrAsnLysLeuGlyLysGlnAlaIleLysIle 229
 ||| ||| ||| |||||
 Db 373 ACACGAC 432
 QY 229 AspLysGluArgGlnLysThrGlnGlnMetLysAlaAlaGlnLysSerLysAspLe 249
 ::||| ||| |||
 Db 433 ACACGACGAGACACAGATCAACAGCAACAGCAACAGCTACAGCAGCAGCAGCAAA- 490
 QY 249 uGluGlyThrMetLysPheValAsnThrValMetIleAlaValSerValAlaIleThrVa 269
 ::||| ||| ||| |||||
 Db 491 -CAGCA-ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 548
 QY 269 IleSerIleValAlaAlaIlePheThrCysGlyAlaGlyLeuAlaGlyLeuAlaAlaG 289
 ::||| ||| ||| |||||
 Db 549 AGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 608
 QY 289 yAlaAlaValGlyAlaAlaAlaAlaGlyGlyAlaAlaGlyAlaAlaAlaAlaThrThrVa 309
 ||||| ||||| |||||
 Db 609 AACACACACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 668
 QY 309 lAlaThrGlnIleThrValGlnAla 317
 ::||| ||| |||
 Db 669 AACATCTGAC 693
 RESULT 6
 BG929588 619 bp mRNA linear EST 31-DEC-2001
 LOCUS BG929588
 DEFINITION eUSHEST0008 Eth11 Eimeria tenella CDNA clone etsha017 5', mRNA
 sequence.
 ACCESSION BG929588
 VERSION BG929588.1 GI:18002978
 KEYWORDS EST.
 SOURCE Eimeria tenella.
 ORGANISM Eimeria tenella

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: January 27, 2003, 14:10:57 ; Search time 1691 Seconds
(without alignments)
4750.421 Million cell updates/sec

Title: US-09-889-314-2
Perfect score: 2369
Sequence: 1 DTNMISSSGPDKNKNIMS.....LKAVATISGATAGAKHTNNE 496

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_p2n_model -DEV=xlh
-O=/cgn2_1/USPTO-spool/US09889314/unaltd_24012003_144540_10121/app_query.fasta_1.647
-DB=EST -QFMT=fastap -SUFFIX=first -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNIT=bits -STRRT=1 -END=1 -MATRIX=Dlosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR.SCORE=pct -THR.MAX=100 -THR.MIN=0 -ALIGN=45 -MODE=LOCAL
-OUTFMT=pro NORM-ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09889314@cgn2_1.1.899@runat.24012003_144540_10121 -NCPV=6 -ICPV=3
-NO_XLPHY -NO_MAP -LARGEQUERY -NEG.SCORES=0 -WAIT -LONGLOG -DEV.TIMEOUT=120
-NARN.TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6 -Fgapext=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:*
1: em.estba:*
2: em.esthum:*
3: em.estin:*
4: em.estnu:*
5: em.estov:*
6: em.estpl:*
7: em.estro:*
8: em.htc:*
9: gb.est1:*
10: gb.est2:*
11: gb.htc:*
12: gb.est3:*
13: gb.est4:*
14: gb.est5:*
15: em.estfun:*
16: em.estom:*
17: gb.gss:*
18: em.gss.hum:*
19: em.gss.inv:*
20: em.gss.pln:*
21: em.gss.vtl:*
22: em.gss.fun:*
23: em.gss.mam:*
24: em.gss.mus:*
25: em.gss.other:*
26: em.gss.pro:*
27: em.gss.tod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	161.5	6.8	936	17	CNS07CJY
2	155.5	6.6	1150	17	CNS079ZS
3	153.5	6.5	935	17	CNS033D4
4	152.5	6.4	935	17	CNS033D4
5	148	6.2	711	13	BM588321
6	147.5	6.2	619	13	BM588321
7	147.5	6.2	1094	17	CNS076CM
8	144.5	6.1	2295	11	BC016304
9	142	6.0	938	17	CNS07CG2
10	142	6.0	2524	11	BC028236
11	140	5.9	1845	11	AY103583
12	140	5.9	4272	11	BC016144
13	137.5	5.8	746	17	A2624776
14	136.5	5.8	987	11	AK013560
15	136	5.7	1194	9	AL334303
16	135.5	5.7	3344	11	BC028681
17	135	5.7	2589	11	BC028951
18	134.5	5.7	485	9	AA930444
19	134.5	5.7	809	17	CNS02310
20	134.5	5.6	1123	17	CNS06RUI
21	132.5	5.6	1027	11	AY107616
22	132	5.6	888	17	CNS075P4
23	131	5.5	892	17	CNS075P4
24	130.5	5.5	1101	17	CNS05HL2
25	130	5.5	887	14	BO735450
26	129	5.4	1745	11	BC007388
27	128.5	5.4	443	17	FR0008252
28	128.5	5.4	751	12	BG656282
29	128.5	5.4	1000	13	BM556446
30	128	5.4	619	17	FR0006944
31	128	5.4	711	13	BM588321
32	128	5.4	773	13	BE585195
33	127.5	5.4	494	10	AW660984
34	127.5	5.4	805	12	BG676639
35	127	5.4	644	14	BQ838510
36	127	5.4	785	12	BG246146
37	126.5	5.3	1590	11	AK014874
38	126.5	5.3	1792	11	AY103948
39	126.5	5.3	2477	11	AK014336
40	126	5.3	786	13	BI106945
41	126	5.3	3110	11	AK019964
42	126	5.3	3441	11	BC030898
43	126	5.3	4088	11	BC021514
44	125	5.3	526	17	AZ312601
45	125	5.3	559	17	A2661219

ALIGNMENTS

RESULT 1
CNS07CJY
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS

CNS07CJY 936 bp DNA linear GSS 08-JUL-2001
T7 end of clone BD0A003E07 of library BD0A from strain CBS 94 of
Candida tropicalis, genomic survey sequence.
AL339268
AL339268.1 GI:12222681
GSS
Candida tropicalis.
Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
Saccharomycetales; mitosporic Saccharomycetales; Candida.
1 (bases 1 to 936)
Soulet,J.L., Aigle,M., Artiguenave,F., Blandin,G.,
Boulotin-Fukuhara,M., Bon,E., Brothier,P., Casaregola,S.,